ORIE Nutrition Research in Northern Nigeria

Quantitative Impact Evaluation of the WINNN Programme

Volume 2: Technical Companion and Methodological Annexes

Operations Research and Impact Evaluation

Paul Jasper, Paola Vargas, Julia Hug, Aly Visram, Adiba Khaled and Patrick Ward

August 2017

Acknowledgements

Operations Research and Impact Evaluation (ORIE) is led by Oxford Policy Management (OPM) in conjunction with three other UK-based institutions, the London School of Hygiene and Tropical Medicine (LSHTM), the Institute of Development Studies (IDS) and Save the Children UK (SCUK), and four Nigerian partners, the University of Ibadan, Kaduna Polytechnic, Ahmadu Bello University at Zaria (ABU), and the Food Basket Foundation International (FBFI).

ORIE is funded by the Department for International Development of the UK Government and implemented in collaboration with the Government of Nigeria.

This report is the culmination of the efforts of many teams over the five years from 2012 to 2017. The team leaders were: Elaine Ferguson (operations research); Aly Visram (quantitative impact evaluation); Emma Jones (qualitative impact evaluation); Sarah Keen (economic evaluation); Frances Hansford (gender); and Tom Barker and Marta Moratti (evidence dissemination and uptake). Many other Nigeria- and UK-based team members contributed. The international team leader was Andrew Tomkins, the project director was Patrick Ward and the national team leader was Vincent Ahonsi. The team members who contributed directly to the preparation of this report were: Paul Jasper, Aly Visram, Julia Hug, Adiba Khaled, Emma Jones, Patrick Ward, Andrew Tomkins, Paola Vargas, Michele Binci, Frances Hansford and Mehroosh Tak.

The many insightful contributions from members of the Working to Improve Nutrition in Northern Nigeria (WINNN) implementing partners, the Department for International Development (DFID) Nigeria office, and various federal and state government agencies were crucial in ensuring that our findings are relevant and useful to a wide set of Nigerian stakeholders, as well as international audiences more broadly. The work that lies behind this report could not have been realised without ongoing and dedicated support from many sources, including leadership and coordination from the ORIE National Team Leader Vincent Ahonsi, administrative support from the ORIE Nigeria office Adesoye Aro, and a number of project officers in OPM's Project Support Unit including most recently Laura Shelton and Carina Dale.

Most importantly, we are extremely grateful to all the many people in the WINNN-supported states – including programme beneficiaries, health workers, community volunteers, community leaders, members of civil society organisations (CSOs) and community-based organisations (CBOs), and government officials – who have generously shared their time, experiences and insights with our research teams over the last five years. Without them this report, and the studies which underlie it, would not have been possible.

Suggested citation: Jasper, P., Vargas, P., Hug, J., Visram, A., Khaled, A., Ward, P. (2017), 'Quantitative Impact Evaluation of the WINNN Programme – Volume 2: Operations Research and Impact Evaluation', Oxford Policy Management, Oxford, UK

This research was carried out by the ORIE consortium. The ORIE project is managed by Patrick Ward at OPM. For further information on this report, please email psu.ORIE@opml.co.uk or see the website: http://www.heart-resources.org/tag/orie/
The contact point for the client is Melkamnesh Alemu: m-alemu@dfid.gov.uk

ORIE

Oxford Policy Management
Level 3, Clarendon House
Cornmarket Street
Oxford OX1 3HJ
United Kingdom

Tel +44 (0) 1865 207 300
Fax +44 (0) 1865 207 301
Email admin@opml.co.uk
Website www.opml.co.uk

Table of contents

List of tables and figures

Abbreviations

ANC	Antenatal care
ATT	Average treatment effects on the treated
CAPI	Computer-assisted personal interview
CMAM	Community-based management of acute malnutrition
C	Control
DFID	UK Department for International Development
DID	Difference-in-difference
EA	Enumeration area
ENA	Emergency Nutrition Assessment
EQuALS	Evaluation Quality Assurance and Learning Services
FANTA	Food and Nutrition Technical Assistance
FAO	Food and Agriculture Organization
GAM	Global acute malnutrition
HAZ	Height-for-age Z-score
ICC	Intra-cluster correlation
INGO	International non-governmental organisation
ITP	Inpatient Therapeutic Programme
ITT	Intention to treat
IYCF	Infant and young child feeding
JMP	Joint Monitoring Programme
Km	Kilometre
LAZ	Length-for-age Z-score
LGA	Local Government Area
M&E	Monitoring and evaluation
MAM	Moderate acute malnutrition
MDG	Millennium Development Goal
MICS	Multiple Indicator Cluster Survey
MNCH	Maternal, newborn and child health

MUAC	Mid-upper arm circumference
NDHS	Nigeria Demographic and Health Survey
NN	Nearest neighbour
NNHS	National Nutrition and Health Survey
OPM	Oxford Policy Management
ORIE	Operations Research and Impact Evaluation
ORS	Oral rehydration solution
OTP	Outpatient Therapeutic Programme
PRRINN-MNCH	Partnership for Reviving Routine Immunisation in Northern Nigeria, Maternal Newborn and Child Health Initiative
PSM	Propensity score matching
PSU	Primary sampling unit
SAM	Severe acute malnutrition
SMART	Standardized Monitoring and Assessment of Relief and Transitions
SSS	Salt sugar solution
TORs	Terms of reference
T	Treatment
UNHCR	United Nations High Commissioner for Refugees
UNICEF	United Nations International Children's Emergency Fund
WAZ	Weight–for-age Z-score
WHO	World Health Organization
WHZ	Weight-for-height Z-score
WINNN	Working to Improve Nutrition in Northern Nigeria

1 Introduction

1.1 Objective of this report

This report has four main objectives. First, it aims to be a technical and methodological companion to Volume 1 of the quantitative impact evaluation report (Quantitative Impact Evaluation of the WINNN Programme – Volume 1: Operations Research and Impact Evaluation, 2017), where our main findings are presented. Thus, the methodology used in the evaluation is discussed here in detail, including explanations of the theoretical models and robustness checks implemented. Second, it gives additional background information about the process of defining the scope and implementing our evaluation that are not included in Volume 1. Third, it presents additional results that are not included in Volume 1. Finally, this report responds to any additional requirements posed by the UK Department for International Development's (DFID's) internal quality assessment system (EQuALS). The target audience for this report are those key stakeholders who are interested in the technical and methodological details of the evaluation. For a discussion of the key findings, please refer to Volume 1.

1.2 Structure of this report

The remainder of this report is structured as follows: Section 2 provides additional background to this evaluation, with a particular focus on the context of the Working to Improve Nutrition in Northern Nigeria (WINNN) programme and this evaluation (2.1), the involvement of stakeholders throughout the process of this evaluation (2.2), the dissemination plans (2.3), the ethical considerations involved in this evaluation (2.4), a declaration regarding conflicts of interest within this evaluation (2.5), the departures from the original terms of reference (TORs) (2.6), the evaluation team (2.7), and how this evaluation addresses the Paris Declaration principles (2.8).

Section 3 provides a technical discussion of our quantitative research component. Section 3.1 summarises the data used. This section discusses the methodology behind the sampling approach and sampling weights (3.1.1), and presents our analysis of sample attrition (3.1.3), as well as the key background characteristics of our final sample (3.1.2). It also provides details of how these data were collected (3.1.4) and how the key indicators were created for this analysis (3.1.6).

Section 3.2 presents further results of a data quality assessment of anthropometric data used in this study. Section 3.3 provides an in-depth discussion of the impact evaluation strategy, including the difference-in-difference (DID) methodology and its use in the context of the current evaluation (3.3.1), the results of the various estimation strategies conducted and the possible limitations to the DID analysis and what steps have been taken to address these (3.3.3). Section 0 presents the technicalities of the propensity score matching (PSM) approach used to understand the specific effect of components of some WINNN-supported activities. Section 3.5 provides in-depth details on the supplementary correlation analysis presented in Box 2 of Volume 1, and its rationale.

Section 4 puts the impact evaluation results into a broader context. Section 4.1 discusses the generalisability of our results to Nigeria and outside Nigeria. Section 4.2 compares the findings of the Operations Research and Impact Evaluation (ORIE) impact evaluation to estimates produced by other surveys collecting similar types of information, such as the NDHS survey.

Finally, this report includes annexes that provide information on the original TORs, the Inception Report of this evaluation, the sampling data used in this evaluation (response rates, item non-response, sampling weights, precision of estimates), the data collection instruments, and key definitions of indicators that were constructed as part of this evaluation.

2 Background to the impact evaluation

2.1 Context of the evaluation: malnutrition and malnutrition programming in Nigeria

This section presents key elements of the malnutrition context in northern Nigeria at the time when WINNN was designed, as articulated in the DFID Business Case for the WINNN programme (2011).

2.1.1 The health context

Undernutrition is a leading cause and contributor to deaths in children under five in Nigeria. According to NDHS 2008, at the time of that survey:

- one in four children was underweight and 9% were severely so;
- in northern Nigeria, a third of children under five were underweight, half were stunted, and a fifth were wasted; and
- in the 10 northernmost states, there were an estimated 3.9 million children under five with stunting and 900,000 with severe and acute malnutrition (SAM) (DFID Business Case, 2011).

Table 1 Prevalence of undernutrition in children under five in Africa, sub-Saharan Africa, and Nigeria in 2011

	Africa (%)	Sub-Saharan Africa (%)	Nigeria (%)
Underweight	28	27.1	38.2
Wasting	10	12.4	23.3
Stunting	38	36.8	52.6

Source: DFID Business Case, 2011.

There are also important differences across socioeconomic statuses in terms of being affected by undernutrition, and gender inequality is known to have an important impact on undernutrition, as women are the main caregivers for children and are usually affected by differences in social status.

Undernutrition has important consequences in regard to child development: it is an important contributor to the death of children under five. About 1 million children die in Nigeria each year (DFID, 2011), and, according to the NDHS (2008), the situation is even worse in the north of the country, where the under-five mortality rate was 40% higher than in the rest of the country at the time of the survey. Undernutrition also affects brain development and cognitive abilities, and can contribute to delayed enrolment and poorer learning outcomes at older ages (Omilola *et al*., 2010). This presents a threat to long-term development in Nigeria.

Maternal undernutrition is also likely to be a key contributing factor to child undernutrition and slowed development. Maternal undernutrition and young maternal age are important determinants of low birth weight, which has long-term consequences for child growth. The NDHS (2008) estimated that in the north-east and north-west of Nigeria, 20% of women were undernourished.

2.1.2 The policy context

Globally, improved nutrition links to several of the Millennium Development Goals (MDGs) including MDG 1 (eradicate extreme poverty and hunger), MDG 4 (reduce child mortality) and MDG 5 (improve maternal health). Achieving those goals will involve improving nutrition. Undernutrition has also become an international development priority through the efforts of the Scaling up Nutrition movement.

In Nigeria, the WINNN programme fit within the Nigerian national policy and strategy for nutrition, as set out in the National Policy on Food and Nutrition in Nigeria (2001) and the National Plan of Action on Food and Nutrition (2004). In particular, the policy highlights as a priority the reduction of undernutrition, with a focus on SAM for children under five, the 'enhancement of care-giving capacity within households with respect to child feeding and child care practices', and the reduction of micronutrient deficiencies. It also highlights the importance of raising understanding of malnutrition in Nigeria at all levels, especially with respect to its causes and solutions (National Plan of Action on Nutrition, 2004).

2.1.3 Existing nutrition programmes in Nigeria

Federal and state-level government bodies have been established in Nigeria with mandates relating to nutrition, such as the National Committee on Food and Nutrition and the Nutrition Partners Group.

Additionally, many donors, multilateral organisations and international non-governmental organisations (INGOs) are active in Nigeria and provide direct interventions to address undernutrition in Nigeria. The United Nations Children's Fund (UNICEF) has nutrition programmes across nine states in northern Nigeria, and provides support for infant and young child feeding (IYCF) practices, micronutrient deficiency and treatment of SAM. UNICEF's model of intervention is to strengthen the host country's capacity. Save the Children UK runs a prevention, detection and treatment programme for acute malnutrition, while Action Against Hunger are providing health services with UNICEF, government partners and the DFID-funded Partnership for Reviving Routine Immunisation in Northern Nigeria, Maternal Newborn and Child Health Initiative (PRRINN-MNCH).

The United States Agency for International Development provides comprehensive childhood survival programmes, which include elements of nutrition, such as breastfeeding and micronutrients. The World Health Organization (WHO) provides technical support to the Federal Ministry of Health and to states to develop nutrition policies.

Additionally, indirect interventions are also in place that target basic and underlying causes of undernutrition. These are in part initiatives already funded by DFID. Among the causes of malnutrition currently addressed by DFID programming are:

1. infectious disease, immunisation and access to health services, addressed by DFID through, for example, PRRINN-MNCH and the Partnership to Transform Health Systems (PATHS2);

2. IYCF practices, also addressed partly through PRRINN-MNCH;

3. inadequate access to water and sanitation facilities, addressed through a partnership between DFID, UNICEF, and the Nigerian government using local and community-driven interventions to improve water and sanitation provision; and

4. food security and poverty, addressed through for example, the 'Promoting Pro-Poor Opportunities through Commodity and Service Markets' programme (Procom) or the Growth and Employment in States programme.

There are two main issues surrounding nutrition programming in northern Nigeria, as recognised by the DFID Business Case (2011). First, there are still limited evidence-based direct interventions. Second, as seen above, many programmes intervene indirectly to address the causes of undernutrition in northern Nigeria but there is a little coordination of these programmes around the issue of nutrition, and their impact on nutrition is still uncertain and also requires more evidence.

The WINNN programme was therefore designed to address some of these gaps by:

1. delivering evidence-based direct interventions;
2. promoting linkages between programmes relevant to nutrition; and
3. enhancing understanding of undernutrition in northern Nigeria through operational research, including this impact evaluation.

2.2 Stakeholders' involvement

All stakeholders were involved in the design of the evaluation and throughout the evaluation, including donors, implementers, the Nigerian Federal Ministry of Health and the National Primary Health Care Development Agency.

During ORIE's inception phase, and in order to understand the organisations active in the northern Nigerian nutrition landscape and the pathways between them, a map of the different actors active in this sector and their influence was drawn up by the ORIE team in stakeholder workshops. This also made it possible to understand the barriers and opportunities to the uptake and use of research evidence in that context. Important messages emerging from both the workshops and the wider stakeholder engagement have been included within the design of this evaluation.

Additionally, both WINNN and the donor counterpart (DFID) were consulted in relation key design features of the quantitative evaluation. For example, the Inception Report (see Annex B) explicitly states that 'due to the nature of the WINNN Programme, the Treatment Group [was] identified in collaboration with Government stakeholders and DFID'.

Throughout the evaluation process, there has been a constant and robust process of reviewing and updating all reports relating to this evaluation, including the present one. Implementation partners received full drafts of all reports for comments, and face-to-face presentations of methodologies employed and of findings that resulted from evaluation analyses. Comments were then addressed and, in a second round, a similar iterative process took place together with DFID. Finally, internal and external quality assurance processes were implemented before reports were finalised to ensure the robustness and adequacy of the methodology, as well as making sure that these are useful and clear products that can inform policy and decision-making in the future.

All key evaluation publications are published online and are publicly accessible, presented in person in multi-stakeholder events, and have easy-to-read summaries.

Note that this impact evaluation is just one component of the larger ORIE work, through which other evaluations are being produced (qualitative, costing, and cost-effectiveness). These other products have also received regular input and feedback from both implementer and donor counterparts.

2.3 Dissemination plan

All key stakeholders at federal and state levels have been invited to regular dissemination workshops and events throughout the course of the evaluation. All major products will be disseminated in a federal workshop and state-level dissemination events. To improve take-up, the dissemination strategy and events have been coordinated as joint efforts with the implementation agencies.

Baseline findings were widely shared so that the programme could make adjustments in line with those findings. Equally, the endline analysis is being prepared on a timeline to coincide with the end of the WINNN programme. Our dissemination strategy is timed such that findings, lessons, and recommendations can be used in any continuation, scale-up, or government takeover of the programme.

The report and its findings will be published in a full report, a summary report and PowerPoint presentations. In addition, separate thematic summaries of findings are being prepared and tailored to the needs and different interests of the various stakeholders involved in this evaluation.

2.4 Ethical considerations

Ethical approval was obtained before carrying out the study after review by the National Health Research Ethics Committee sitting in the Federal Ministry of Health of Nigeria.

Consent in local language was sought from all respondents before any questionnaire was conducted for data collection (see Annex D.1 for the consent forms) and interviewers were trained in intensive training programmes before each round of data collection to ensure that data were collected in way that was respectful to the interviewed households and to the local cultural context.

Data will be anonymised before being published. In addition, any severely acutely malnourished children that were identified during household visits were appropriately referred for treatment.

2.5 Conflicts of interest

Finally, it is important to state that the evaluators act independently from the programmes we are evaluating and disclose any potential conflicts of interest that might jeopardise the integrity of the methodology or the outputs of the research/evaluation, should any arise.

The team was able to work freely and independently while ensuring regular feedback from, and the involvement of, various stakeholders. Information sources and their contributions were fully independent of other parties with an interest in the evaluation.

2.6 Departure from TORs

The TORs for this evaluation were drafted in 2011, at a time when the WINNN programme was still being refined and two years before the beginning of this evaluation (see Annex A for the original TORs). Since the WINNN programme design evolved after the TORs were published, and considering that the TORs were very broad – particularly with regard to the quantitative impact evaluation component – the TORs should not be considered the main reference document for this evaluation.

Instead, the Inception Report which was produced at the beginning of this evaluation constitutes the key document of reference. The Inception Report was written following consultations with

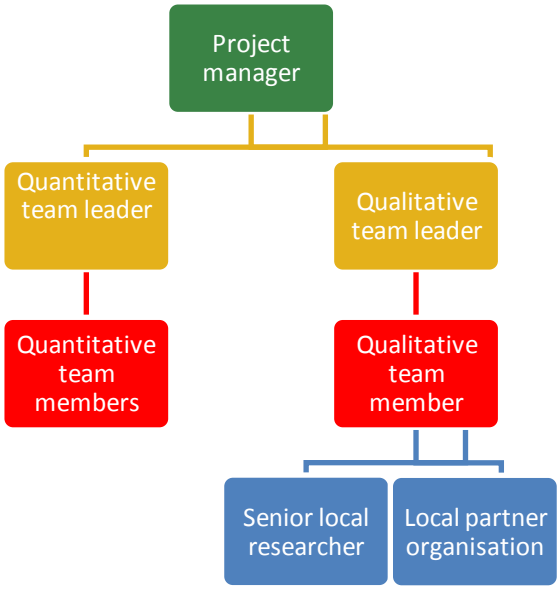
DFID, academics, and Nigerian counterparts to ensure that the goals of this evaluation reflected the interests of different stakeholders. Section 5.3 of the Inception Report Volume I (see Annex B.1 for this full section from the Inception Report Volume I) and the section on the quantitative evaluation component of ORIE of the Inception Report Volume II (see Annex B.2 for the full section from the Inception Report Volume II) define the WINNN programme evaluation requirements and objectives.

This evaluation departs in four minor ways from the plan delineated in the Inception Report. First, while Volume II of the Inception Report outlines the potential possibility of using routine and secondary data to produce an interrupted time series analysis, this was not implemented in the end due to limited routine data available to the team and time constraints. Second, the Inception Reports mentions that the evaluation team had intended for this evaluation to link back to WINNN's monitoring and evaluation (M&E) system to be able to better understand impact or non-impact. The lack of consistent and reliable monitoring data available to the evaluation team meant that it was impossible to undertake this exercise as part of the current evaluation. Third, this evaluation did not report on trends in the mortality of children under five using secondary data in the WINNN Local Government Areas (LGAs), as specified in the Inception Report. Finally, while it was considered during the inception phase, the evaluation sampling design did not, in the end, over-sample children under 12 months, due to budget and time constraints.

2.7 Evaluation team and management

Our management and team structure throughout the evaluation has closely followed what had been suggested in the TORs, our proposal, and the Inception Report. The project was led and managed by Oxford Policy Management (OPM).

Figure 1 Evaluation team and management structure of ORIE



The project was led by Patrick Ward. The project was managed by Aly Visram, who managed the overall ORIE project and this quantitative evaluation. Both ensured the smooth running of administrative issues and communication flows within the team and with respect to stakeholders, including DFID. It is important to emphasise here again that this evaluation was managed in close consultation with stakeholders. This included sharing of findings and recommendations with WINNN and DFID, and taking into account different views and comments from those stakeholders at different stages.

The core technical team for the quantitative component, i.e. this quantitative impact evaluation, was led by Paul Jasper, who provided both overall guidance and technical inputs at baseline and endline. In addition, a pool of OPM analysts has provided technical support to this evaluation.

2.8 Paris Declaration principles

This section aims to explain how the ORIE quantitative impact evaluation fits within the Paris Declaration principles framework.

Table 2 Paris Declaration principles and how this impact evaluation fits within this framework

Principle	Description	In the context of this impact evaluation
Ownership	Developing countries set their own development strategies, improve their institutions and tackle corruption.	National and state-level policy-makers have been engaged throughout the course of the evaluation through dissemination events, meetings and the production of summary publications to increase the use of evidence by policy-makers.
Alignment	Donor countries and organisations bring their support into line with these strategies and use local systems.	This evaluation fits within the national framework and priorities since the National Plan of Action on Nutrition (2004) highlights the importance of raising understanding of malnutrition in Nigeria at all levels, especially with respect to its causes and solutions.
		The WINNN programme itself is in line with Nigeria's nutrition strategies and aims to use local systems to improve delivery of health services for children and women in northern Nigeria.
Harmonisation	Donor countries and organisations coordinate their actions, simplify procedures and share information to avoid duplication.	All stakeholders were involved in the design of the evaluation and throughout the evaluation, including DFID, WINNN, implementers, the Nigerian Federal Ministry of Health and the National Primary Health Care Development Agency.
		All ORIE reports (including this impact evaluation) are available online and anonymised datasets will be made publicly available at the end of the evaluation in the World Bank Microdata Library (http://microdata.worldbank.org/index.php/home).
Mutual accountability	Donors and developing countries are accountable for development results.	Throughout the evaluation process, there has been a constant and solid process for reviewing and updating all reports relating to this evaluation, including the present one. Implementation partners received full drafts of all reports for comments, and face-to-face presentations of methodologies employed and of findings that resulted from evaluation analyses. Comments were then addressed and, in a second round, a similar iterative process took place together with DFID. Finally, internal and external quality assurance processes were implemented before reports were finalised.
		All key stakeholders at federal and state levels have also been invited to regular dissemination workshops and events throughout the course of the evaluation.

3 Quantitative analysis

3.1 Data

3.1.1 The survey

This study focuses on children within the age range of 0 to 35 months, on the mothers of those children who are of reproductive age (that is between 15 and 49 years), and their households. The data used in this quantitative impact evaluation are survey data collected on such children and mothers in a panel of households between baseline and endline. The purpose of this section is to provide further details on how this survey was designed and how data were collected both at baseline and endline.

Sampling strategy

The sampling strategy used to collect data for this study was a multistage random sampling method. It is important to mention here that, as explained in Volume 1, treatment status in this impact evaluation was not allocated randomly to LGAs. Rather, state officials and WINNN implementing partners selected certain LGAs for treatment. Hence, control LGAs were also not chosen at random, but matched to the treatment LGAs based on a list of observed characteristics (see Annex B for relevant sections on the selection of control groups in the Inception Report). In each state, three treatment and three control LGAs were selected for this study. In total, 24 LGAs were chosen for this study.

Within this total number of 24 LGAs, the primary sampling units (PSUs) were enumeration areas (EA) as defined by the 2006 Nigerian Census. In the context of this evaluation, these EAs are referred to as 'communities'. A sampling frame was not available at the outset of the evaluation. Hence, a list of all EAs across treatment and control LGAs was prepared. Thirty-five EAs were then sampled per LGA using a random draw, with each EA in the same LGA having the same probability of being selected. In total, 840 EAs were sampled for this study. This is the first stage of sampling.

The second sampling stage occurred within each EA, where households with at least one child under the age of three (i.e. between 0 and 35 months) were sampled. The definition of the household used was 'a person or group of related or un-related persons that live together in the same dwelling unit and acknowledge one adult male or female as the head of the household' (see Annex E for more details on this definition). These secondary sampling units were selected because it is assumed that the WINNN interventions, aimed at infants and young children, are most likely to affect children under the age of three. Hence, the aim was to obtain a sample in which there would be a high number of children within this age range. Where a sampled EA did not contain at least 10 households with at least one child under the age of three, an adjacent EA was linked to it.

Since no sampling frame existed, a listing exercise was conducted within each EA, during which a census of all households was collected. Detailed maps of the EA were prepared to enable the baseline and follow-up survey teams to identify sampled households. Listing was implemented using brief interviews aimed at collecting basic household information to identify households as well as the nature of their composition.

Due to the brevity of the listing interview, only an estimated number of children under the age of three could be collected for each household. Such information may often be unreliable due to the

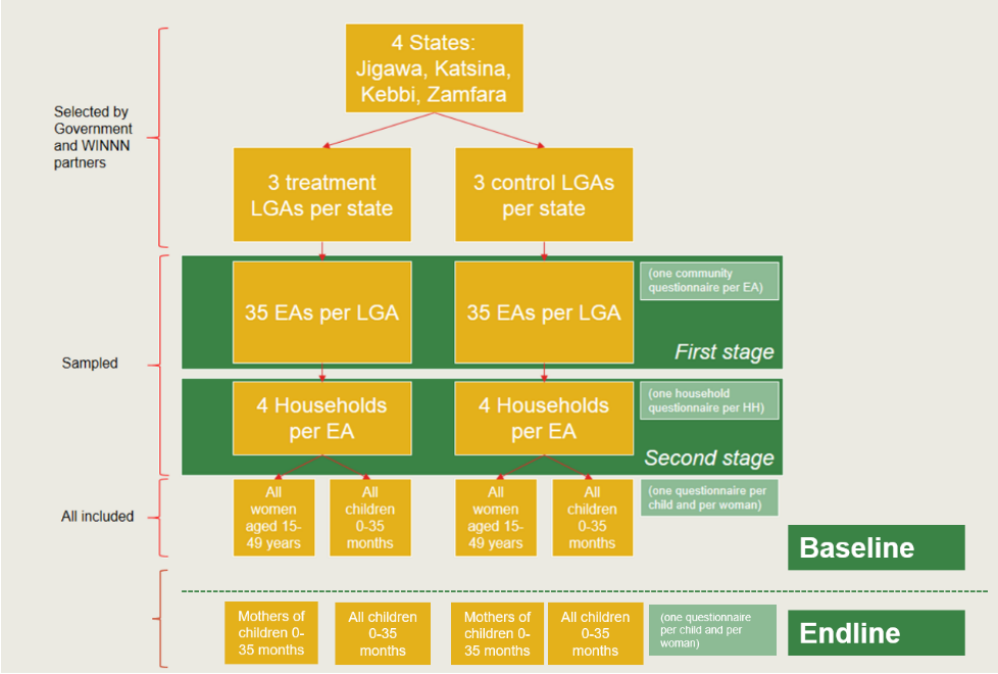
large size of households and non-accurate perceptions of the age of children by respondents. Hence, a sample unit replacement protocol was implemented at the level of the EA.

A random sample of households was independently drawn and provided to the baseline survey team. In total, four households were sampled per EA, thus making a total of 3,360 households throughout the study (four households per EA x35 EAs per LGA x six LGAs per state x four states). In addition to the sampled households, the baseline survey team was also provided with a randomised list of potential replacement households within each EA, contained in sealed envelopes by the baseline team supervisor.

Overall, this evaluation aimed to sample at baseline 3,360 households (1,680 treatment, 1,680 control), in 840 EAs (220 treatment, 220 control), in six LGAs (three treatment, three control), in four states. Figure 2 below outlines the different steps of the sampling strategy.

Since this evaluation uses a panel of households, at endline enumerators were asked to track down households that had been interviewed at baseline to be interviewed again. However, in the interest of time, only households with children under three and their mothers were interviewed at endline, i.e. interviews were not conducted in all households that were approached at baseline but rather only households that had children within the age range of interest for this evaluation. Note that this automatically leads to attrition in our sample between baseline and endline, because not all households in which children under three could be found at endline also included children under the age of three at endline. We deal with attrition in Section 3.1.3 of this report. Figure 2 below outlines the difference between baseline and endline.

Figure 2 Sampling strategy



Source: Adapted from ORIE baseline report

It is important to emphasise again that this sampling ensured that our estimates are representative of the population of households with children aged 0–35 months in treatment and control LGAs overall at baseline – i.e. across the geographical area that these LGAs cover – irrespective of the exact location of WINNN-supported activities within treatment LGAs. This means that, within LGAs, certain indicator estimates presented in Volumes I and II might vary geographically and could be

higher in catchment areas around facilities or locations where WINNN has focused its efforts within treatment LGAs. Our estimates, in contrast, are averages that hold across treatment LGAs overall.

Sample size

Sample size and power calculations have been computed at baseline for the intention to treat (ITT) impact estimator and for some key outcome variables, including the prevalence of wasting, stunting and underweight among children. As a result, the baseline sample size of children under three was 3,463 in the treatment group and 3,370 in the control group, making a total of 6,833 children aged 0–35 months in the study. Across 840 clusters (EAs) with an intra-cluster correlation (ICC) of 0.09 for stunting, 0.13 for underweight and 0.02 for wasting, this sample size was deemed to be sufficient to detect a decrease in stunting by five percentage points (from 58% to 53%), in underweight by five percentage points (from 41% to 36%), and in wasting by three percentage points (from 16% to 13%). More details on the calculations carried out at baseline on power, the design effect, and the minimum detectable effect can be found in Annex C.1).

At endline, survey teams aimed to track down all clusters and households that had been interviewed at baseline. However, only 829 out of 840 clusters from baseline were visited at endline due to access issues. In addition, some households could not be found due to security issues and migration (mostly cattle rustling in Katsina and Jigawa, and migration of Fulani settlements in Zamfara). An overview of the final sample size at baseline and endline is presented in Table 3 below.

Table 4 presents a breakdown of this sample by state.

The results presented in this table show that at endline 2,722 households with eligible children were visited and individuals interviewed there, which represents a decrease of about 19% compared to the baseline figure of 3,355.[1] Note that this translates into roughly similar proportional decreases in sample sizes for children and mothers (15%). Section 3.1.3 delineates how this attrition was dealt with at the analysis stage in order to tackle any potential for bias. Note in particular that we calculate point estimates of outcomes using attrition-corrected weights, and that these imply that some households are dropped from our final analysis (households with no values in covariates used in the attrition model will not have attrition-corrected weights assigned). This means that most indicators presented in this evaluation will be estimated on a smaller sample size than the one reported in Table 3 and

[1] This was calculated as follows: 1 – (3,355/2,722) = 18.88

Table 4 below. Please note that Section 3.1.2 provides additional evidence that attrition did not lead to substantive differential changes across treatment and control areas with respect to the background characteristics of individuals included in our analysis.

Table 3 Final sample size

Level of analysis	Baseline	Endline	Baseline		Endline	
			Treatment	Control	Treatment	Control
Interviewed households	3,457	3,229	1,710	1,747	1,595	1,634
Eligible households (with eligible children 0–35 months)	3,355	2,722	1,677	1,678	1,347	1,375
Children 0–35 months	6,833	5,567	3,463	3,370	2,777	2,790
Mothers of children 0–35 months (mothers of reproductive age 15–49 years)	5,708	4,784	2,855	2,853	2,406	2,378
Communities	840	829	420	420	411	418

Table 4 Final sample size by state

Level of analysis	State	Baseline	Endline	Baseline		Endline	
				Treatment	Control	Treatment	Control
Households (with eligible children 0–35 months)	Jigawa	839	682	419	420	327	355
	Katsina	837	688	418	419	345	343
	Kebbi	840	698	420	420	349	349
	Zamfara	839	654	420	419	326	328
Children 0–35 months	Jigawa	1,652	1,413	771	881	633	780
	Katsina	1,590	1,338	813	777	692	646
	Kebbi	1,879	1,473	997	882	780	693
	Zamfara	1,712	1,343	882	830	672	671
Mothers of children 0–35 months (mothers of reproductive age 15-49 years)	Jigawa	1,398	1,184	645	753	532	652
	Katsina	1,293	1,134	645	648	586	548
	Kebbi	1,572	1,298	823	749	697	601
	Zamfara	1,445	1,168	742	703	591	577
Communities	Jigawa	210	210	105	105	105	105
	Katsina	210	207	105	105	102	105
	Kebbi	210	210	105	105	105	104
	Zamfara	210	201	105	105	99	104

Note that, at baseline, not all the households initially surveyed were included in the final analysis. About 3% of the total initially surveyed (N = 102) were either replaced during data collection or dropped during data cleaning. This was due to two reasons: first, enumerators identified some households as not eligible for the survey as no children under the age of three could be identified at the time of the interview. In this case, a replacement household from the same EA was selected from the replacement data.

In addition, both at baseline and endline, at the stage of data cleaning, a cleaned child age variable identified some instances where households did not actually include a child under the age of three. Often, cleaning of the age variable revealed that one child in the household was just over the three-year cut-off by a matter of months, and thus was not eligible for the study. These households were then dropped from the analysis. We do not expect this to have any effect on our overall design as the baseline survey achieved 99.9% of the expected sample. Annex C.2 on the response rates, which highlights some of the differences between the surveyed sample and the sample included in the analysis.

Sampling weights and structure of the data

All results presented in Volume 1 and Volume 2 of this report are inclusive of sampling weights and take into account the survey structure of the data. Annex C.4 explains how the sampling weights were calculated for our analysis and how this was taken into account.

3.1.2 Sample background characteristics

Characteristics of households and communities

The following section describes the general characteristics of the communities (840 at baseline and 829 at endline) and households (3,457 at baseline and 3,229 at endline) surveyed in the study. The community questionnaires were generally answered by two to three prominent members of the community that were selected by the emir of each EA, and who typically were health professionals, businessmen, government officials, or civil servants. Hence, responses to this

questionnaire mainly reflect the knowledge and views of a relatively small sample of educated, powerful and largely male informants from the community. It is also important to reiterate that all distances in the community questionnaire use the emir's palace as the point of reference, as this was typically situated in the centre of the community. The household questionnaire was answered by the head of the household – or the person acting as head of household in the absence of the head of household.

Table 5 reports the occurrence of natural disasters, drought and flood, in the communities during the last 12 months. 34% of community informants reported their community experiencing a drought in the last year at baseline, and 33% reported this at endline. While at baseline treatment areas were reported to be significantly more affected by floods than control areas (13% more), this was no longer the case at endline. Overall, communities experienced fewer floods at endline than at baseline.

Table 5 Community indicators

Indicator name	Baseline				Endline			
	Total	T	C	Diff (T-C)	Total	T	C	Diff (T-C)
Percentage of communities...								
That have experienced drought in the last 12 months	33.9	30.9	36.4	-5.5	33.3	29.7	36.4	-6.7
N	*818*	*411*	*407*		*787*	*395*	*392*	
That have experienced flood in the last 12 months	53.2	60.2	47.3	12.9***	45.5	48.0	43.4	4.6
N	*819*	*412*	*407*		*674*	*334*	*340*	

Notes: The 'N' shows the number of unweighted observations. Significance levels of the T-C difference and DID are reported with stars: ***Significant at 99.9% level, **Significant at 99% level, *Significant at 95% level.
T: Treatment
C:Control

Table 6 presents descriptive statistics on household demographics. Household size remained stable from baseline to endline, with, on average, 12 individuals per household. Overall, the demographic composition of a household remained the same between baseline and endline, with an average of five adults (18 years +), seven minors (0–17 years old), two very young children (0–3 years old), three women of reproductive age and less than one elderly household member (aged 65 and above) per household.

The demographic dependency ratio measures the number of individuals per household typically in the labour force and those typically not in the labour force (i.e. children and the elderly). Hence, it is a rough indicator of the economic burden on household members who are able to work. As the ratio increases, the more the household is reliant on able-bodied members. The demographic dependency ratio went from 146 at baseline to 153 at endline. This means that on average there were 1.5 dependents for every working-aged person in a household at both baseline and endline. As there are few elderly household members, the ratio is largely driven by the number of children (0–14 years old) in the household.

Even though nearly all households still have a male head of household (99%), the percentage of households headed by a female has increased from 1% at baseline to 2% at endline. Child-headed households, defined as those with a household head younger than 18, are rare. The average age of the household head increased from 49 at baseline to 51 at endline. Only a third of household heads had any form of formal schooling at both survey points. Even though Islamia education is much more common, our surveys estimate a decrease in the proportion of household heads having completed some Islamia education from 84% at baseline to 74% at endline. Finally, while almost all male household heads were in a marriage (99% at both survey points), the number of

male heads reporting being married in a monogamous relationship decreased from 56% at baseline to 51% at endline, and the number of male heads in a polygamous marriage increased from 43% at baseline to 47% at endline.

From an impact evaluation perspective, it is important to emphasise that these results show that between baseline and endline no significant differences between households in treatment and control areas appear that could bias our results. In effect, the table below shows that our sample is quite balanced at baseline and at endline, when it comes to household demographics across treatment and control groups.

Table 6 Household demographics

Indicator name	Baseline				Endline			
	Total	T	C	Diff (T-C)	Total	T	C	(T-C)
Household								
Mean number of children under three years (0–35 months)	2.0	2.0	2.0	0.0	1.9	2.0	1.9	0.1
Mean number of children (0–12 years)	6.0	6.0	6.1	0.0	6.1	6.2	6.1	0.1
Mean number of minors (0–17 years)	7.1	7.0	7.1	-0.1	7.3	7.3	7.2	0.1
Mean number of adults (18+ years)	5.0	5.0	5.0	0.0	5.0	5.0	4.9	0.1
Mean number of elderly (65+ years)	0.3	0.3	0.3	0.0	0.4	0.4	0.4	0.0
Mean number of women of reproductive age (15–49 years)	2.5	2.5	2.5	0.0	2.6	2.7	2.6	0.1
Mean household size	12.1	12.1	12.2	-0.1	12.3	12.4	12.2	0.2
N	*3,355*	*1,677*	*1,678*		*2,599*	*1,298*	*1,301*	
Mean dependency ratio 1/	145.5	145.9	145.2	0.6	152.7	151.7	153.6	-1.9
N	*3,354*	*1,677*	*1,677*		*2,599*	*1,298*	*1,301*	
Household head								
Type of household head (all households with children 0–35 months)								
Proportion of male-headed households	99.1	99.2	99.0	0.2	98.0	97.9	98.2	-0.2
Proportion of female-headed households	0.9	0.8	1.0	-0.2	2.0	2.1	1.8	0.2
N	*3,315*	*1,657*	*1,658*		*2,599*	*1,298*	*1,301*	
Proportion of elderly-headed households (65+ years)	16.8	16.7	16.9	-0.2	18.1	17.7	18.5	-0.8
Proportion of child-headed households (under 18 years)	0.1	0.1	0.0	0.1	0.1	0.2	0.0	0.1
N	*3,308*	*1,657*	*1,651*		*2,596*	*1,296*	*1,300*	
Mean age of household head	48.5	48.4	48.7	-0.3	51.2	51.2	51.2	-0.1
N	*3,308*	*1,657*	*1,651*		*2,596*	*1,296*	*1,300*	
Proportion of household heads that have any schooling (non-Islamia)	33.3	33.2	33.3	-0.1	34.4	34.5	34.4	0.0
N	*3,333*	*1,667*	*1,666*		*2,588*	*1,275*	*1,313*	
Proportion of household head has any Islamia education	84.0	82.8	85.1	-2.3	74.8	75.3	74.4	0.8
N	*3,335*	*1,670*	*1,665*		*2,691*	*1,330*	*1,361*	
Household head's education level (non-Islamia)								
No education/nursery	66.9	66.9	66.9	0.0	65.6	65.7	65.6	0.1
Primary	13.8	14.8	12.9	1.9	15.3	15.1	15.5	-0.4
Secondary	19.2	18.3	20.1	-1.9	19.0	19.2	18.9	0.3
N	*3,333*	*1,667*	*1,666*		*2,588*	*1,275*	*1,313*	
Marital status of male household heads								
Married (monogamous)	55.9	57.8	54.2	3.7	51.5	52.4	50.7	1.6

Married (polygamous)	43.3	41.5	44.9	-3.5	47.3	46.4	48.1	-1.7
Divorced/ separated	0.3	0.2	0.3	0.0	0.4	0.4	0.3	0.1
Never married	0.3	0.3	0.2	0.1	0.1	0.1	0.2	-0.1
Widowed	0.3	0.1	0.5	-0.3	0.7	0.7	0.7	0.0
Total	100	100	100		100	100	100	
N	*3,275*	*1,640*	*1,635*		*2,654*	*1,314*	*1,340*	

Notes: The 'N' shows the number of unweighted observations. Significance levels of the T-C difference and DID are reported with stars: ***Significant at 99.9% level, **Significant at 99% level, *Significant at 95% level.
1/ Mean dependency ratio measures the proportion of people in the non-working age group (dependents) to those in the working age group (formula: individuals 0–14 + 65+ / individuals 15–64) X 100)

Table 7 presents household-level food security indicators. Following the Food and Agriculture Organisation (FAO) guidelines, household hunger was measured using a simple 'Household Hunger Score' (Deitchler *et al*., 2011). In the questionnaire, questions were asked about three 'hunger' situations: was there ever no food in the household in the four weeks previous to the survey? Did anybody ever go to sleep hungry in the four weeks preceding the survey? Did anybody ever go for 24 hours without eating in the four weeks preceding the survey? The Household Hunger Score assigns each household one point if it answered these questions with 'rarely' and two points for 'often'. Categories were then constructed for little or no hunger (0–1 points), moderate hunger (2–3 points), and severe hunger (4–6 points). Table 7 shows that at baseline, 82% of households experienced little or no hunger according to this index; this proportion increased to 88% at endline.

Table 7 Household-level food security indicators

Indicator name	Baseline				Endline			
	Total	T	C	Diff (T-C)	Total	T	C	Diff (T-C)
Household hunger scale 1/								
Little or no hunger	82.0	80.7	83.1	-2.5	88.1	87.9	88.3	-0.4
Moderate hunger	16.4	17.0	15.8	1.2	11.1	11.4	10.8	0.6
Severe hunger	1.7	2.3	1.0	1.3**	0.8	0.6	0.9	-0.2
Total	100	100	100		100	100	100	
N	*3,311*	*1,659*	*1,652*		*2,599*	*1,298*	*1,301*	

Notes: The 'N' shows the number of unweighted observations. Significance levels of the T-C difference and DID are reported with stars: ***Significant at 99.9% level, **Significant at 99% level, *Significant at 95% level.
1/ A simple indicator to measure household hunger in food insecure areas developed by FANTA II. Categories constructed for little or no hunger (0–1 points), moderate hunger (2–3 points), and severe hunger (4–6 points)

Safe drinking water was not easily available for households in the study area, as shown in Table 8 . Approximately two-thirds of households (60%) at baseline used an improved source of drinking water, which was slightly reduced, to 54%, at endline. Following WHO/UNICEF Joint Monitoring Programme (JMP) for Water and Sanitation Standards (WHO/UNICEF JMP, 2012), these include piped water into dwelling or yard, public taps, tube wells or boreholes, protected dug wells, protected springs, and rainwater collection. Note that at baseline there was a significant difference between treatment and control areas in terms of access to improved sources of drinking water, which seems to have disappeared at endline.

As with safe drinking water availability, the survey found that only a few (5%) households at both survey points had a safe handwashing place, which refers more explicitly to having both water and soap at a designated handwashing place inside the dwelling and being able to show this place to the enumerator. This was confirmed by direct observation

In addition, the sanitation infrastructure was also found to be poor. According to the old definition of improved sanitation, which only includes flush/pour flush toilet (WHO/UNICEF JMP, 2012), 4% of households at both survey points had access to an improved sanitation facility.

Table 8 Household access to utilities

Indicator name	Baseline				Endline			
	Total	T	C	Diff (T-C)	Total	T	C	Diff (T-C)
Proportion of households that have access to the following basic facilities…								
Improved source of drinking water /1	59.5	53.1	65.4	-12.3***	54.0	52.3	55.5	-3.2
N	*3,355*	*1,677*	*1,678*		*2,599*	*1,298*	*1,301*	
Safe hand washing place /2	4.8	4.9	4.8	0.1	5.4	5.3	5.6	-0.3
N	*3,260*	*1,609*	*1,651*		*2,555*	*1,281*	*1,274*	
Improved toilet facility /2	3.8	4.0	3.6	0.4	4.4	4.3	4.4	-0.2
N	*3,355*	*1,677*	*1,678*		*2,599*	*1,298*	*1,301*	

Notes: The 'N' shows the number of unweighted observations. Significance levels of the T-C difference and DID are reported with stars: ***Significant at 99.9% level, **Significant at 99% level, *Significant at 95% level.
1/ Improved drinking water source includes: (i) piped water into dwelling, (ii) piped water to yard/plot, (iii) public tap or standpipe, (iv) tube well or borehole, (v) protected dug well, (vi) protected spring and (vii) rainwater
2/ Safe handwashing place includes both water and soap for households that have a handwashing area inside the dwelling and were able to show it
3/ improved toilet includes flush/ pour flush toilet only

After broadening the definition of improved sanitation to also include a ventilated improved pit latrine and a pit latrine with slab, the proportion with access to improved sanitation facilities increased to 81% at endline, as seen in Table 9. Note that this is mainly due to changes in the definition of the improved facility.

Table 9 Household access to improved sanitation (new definition)

Indicator name	Endline			
	Total	T	C	Diff (T-C)
Proportion of households that have access to the following basic facilities…				
Improved sanitation including three types /1	81.0	81.3	80.8	0.6
N	*2,599*	*1,298*	*1,301*	

Notes: The 'N' shows the number of unweighted observations. Significance levels of the T-C difference are reported with stars: ***Significant at 99.9% level, **Significant at 99% level, *Significant at 95% level. Data from endline survey only.
1/ improved toilet includes (i) flush toilet, (ii) ventilated improved pit latrine, and (iii) pit latrine with slab

Table 10 shows mothers of children aged 0–35 months were 29 years old on average at endline. Only a small proportion of mothers (5%) were younger than 18 at baseline, compared to an even smaller proportion at endline (2%). At endline, almost all mothers (99%) had a spouse or partner, and a large proportion of them got married before turning 18 (93%), resulting in an average age at first marriage of 15 years, which is comparable to the baseline. The average age at first birth was about two years later, at 17 years. Within their lifetime, a mother gives birth to five children on average. Overall, the mother's characteristics remain similar between baseline and endline, and across treatment and control groups, which again provides evidence for balanced samples between treatment and control groups both at baseline and endline.

Table 10 Maternal characteristics in households with a child 0–35 months

Indicator name	Baseline				Endline			
	Total	T	C	Diff (T-C)	Total	T	C	Diff (T-C)
Mother's age								
Mean age (in years) (*mothers 15–49 years*)	27.7	27.9	27.5	0.4	29.2	29.4	29.0	0.5
Proportion of adolescent mothers (*15–18 years*)	4.5	3.7	5.2	-1.4	2.2	2.2	2.2	0.0
N	*5,670*	*2,835*	*2,835*		*4,556*	*2,314*	*2,242*	
Mother's marital status (*mother 15–49 years*)								
Mean age at first marriage	14.7	14.7	14.7	0.1	15.0	15.0	15.0	0.1
N	*5,597*	*2,800*	*2,797*		*4,484*	*2,280*	*2,204*	
Proportion of mothers that have a spouse or partner	98.5	98.4	98.5	0.0	98.6	98.2	98.9	-0.7
N	*5,665*	*2,831*	*2,834*		*4,784*	*2,406*	*2,378*	
Proportion of mothers married under age 18	94.5	94.6	94.5	0.1	93.1	92.9	93.4	-0.5
N	*5,597*	*2,800*	*2,797*		*4,484*	*2,280*	*2,204*	
Mother's child birth status (*mothers 15–49 years*)								
Mean age at first birth	17.0	17.1	16.8	0.2***	17.2	17.2	17.1	0.1
N	5,502	2,750	2,752		4,430	2,243	2,187	
Number of children given birth.	4.6	4.7	4.6	0.1	5.0	5.0	5.0	0.0
N	*5,670*	*2,835*	*2,835*		*4,556*	*2,314*	*2,242*	
Mother's education (*mothers 15–49 years*)								
Proportion of mothers that have any schooling (non-Islamia)	13.4	12.8	13.9	-1.1	15.4	14.4	16.4	-2.0
N	*5,698*	*2,850*	*2,848*		*4,538*	*2,303*	*2,235*	
Mother's education level (non-Islamia)								
No education/nursery	86.8	87.3	86.3	1.0	85.2	86.1	84.4	1.7
Primary	8.8	9.0	8.6	0.3	10.7	9.9	11.4	-1.5
Secondary	4.4	3.8	5.1	-1.3	4.1	4.0	4.2	-0.2
N	*5,660*	*2,830*	*2,830*		*4,765*	*2,395*	*2,370*	

Notes: The 'N' shows the number of unweighted observations. Significance levels of the T-C difference and DID are reported with stars: ***Significant at 99.9% level, **Significant at 99% level, *Significant at 95% level. Mothers are defined as mothers aged 15–49 of children aged 0–35 months.

Table 11 reports the decision-making power of mothers regarding their child's health at endline. Few mothers (3%) take decisions on their own but still, significantly more mothers in treatments areas (4%) have independent decision-making capability in regard to the health of their child than their counterparts in control (2%) areas. Most mothers (73%) in both treatment and control areas make joint decisions or consult with someone else, while 23% of mothers allow someone else to make decisions on a child's health. Of those who take joint decisions, 98% take decisions with their husbands but slightly fewer do so in treatment areas (98%), as compared to control areas (99%). Across both treatment and control areas, 9% of mothers take joint decisions with their co-wife and 7% with their mother-in-law. 99% of mothers seek permission from their husband/head of household/father to go to a health facility.

Table 11 Mother's health-seeking behaviour

Indicator name	Endline			
	Total	T	C	Diff (T-C)
Mothers (*15–49 years*) makes decision regarding child's health:				
On her own	3.2	4.0	2.4	1.6*
Jointly/in consultation with someone	73.8	73.1	74.4	-1.4
Someone else decides	23.0	22.9	23.1	-0.2
Total	100	100	100	
N	*4,538*	*2,303*	*2,235*	
Mother takes joint decision with (*mother 15–49 years who takes joint decision on child health*): 1/				
Co-wife	9.2	9.3	9.1	0.2
Husband	98.4	97.9	99	-1.1*
Mother-in-law	7.2	7.3	7.2	0.0
Other woman in the household	2.0	2.5	1.6	0.9
Other	6.2	6.5	5.9	0.6
N	*3,390*	*1,731*	*1,659*	
Mother seeks permission to go to health facility (*mother 15–49 years*)	98.7	98.2	99.2	-0.9*
N	*4,436*	*2,236*	*2,200*	

Notes: The 'N' shows the number of unweighted observations. Significance levels of the T-C difference are reported with stars: ***Significant at 99.9% level, **Significant at 99% level, *Significant at 95% level.
Data from endline survey only.
Mothers are defined as mothers aged 15–49 of children aged 0–35 months.
1/ Categories do not add up to 100% because this is multiple choice question.

Table 12 shows that the majority of children surveyed (67%) were between 12 and 35 months old at both baseline and endline, and there were about as many males as there were females. At both survey points, most children were not the first-born children of their mothers and thus they had at least a few older siblings.

Table 12 Child characteristics in households with a child 0–35 months

Indicator name	Baseline				Endline			
	Total	T	C	Diff (T-C)	Total	T	C	DIFF (T-C)
Age (in months) (*children 0–35 months*)								
0 - 5	16.9	16.2	17.5	-1.2	17.4	17.2	17.7	-0.5
6 - 11	16.6	16.4	16.8	-0.4	16.8	17.6	15.9	1.7
12 - 23	31.5	31.6	31.5	0.1	33.2	33.3	33.2	0.1
24 - 35	35.0	35.8	34.2	1.5	32.6	31.9	33.2	-1.3
Total	100	100	100		100	100	100	
N	*6,828*	*3,460*	*3,368*		*5,295*	*2,671*	*2,624*	
Sex (*children 0–35 months*)								
Male	50.3	49.3	51.2	-1.9	50.9	51.0	50.8	0.2
N	*6,826*	*3,460*	*3,366*		*5,302*	*2,677*	*2,625*	
Birth order (*children 0–35 months*)								
1	21.2	20.7	21.7	-0.9	18.1	17.8	18.3	-0.5
2 - 3	38.0	39.1	37.0	2.2	41.9	41.3	42.4	-1.1
4 - 5	27.2	27.7	26.7	1.1	25.9	26.3	25.5	0.8
6+	13.5	12.3	14.6	-2.3	14.1	14.5	13.8	0.8
Total	100	100	100		100	100	100	
N	*6,714*	*3,395*	*3,319*		*5,219*	*2,637*	*2,582*	

Notes: The 'N' shows the number of unweighted observations. Significance levels of the T-C difference and DID are reported with stars: ***Significant at 99.9% level, **Significant at 99% level, *Significant at 95% level.

Table 13 looks at the prevalence of diarrhoea in children aged 0–35 months. 4% more children suffered from diarrhoea in the previous two weeks in treatment areas (40%) than control areas (36%). Of those children, 11% suffered from bloody diarrhoea, 70% sought advice/treatment for their illness from any source, 24% received oral rehydration solution (ORS) and 5% received salt sugar solution (SSS) at a clinic, hospital or health centre. There was no significant difference between treatment and control areas for these indicators, except for SSS treatment, which was slightly higher in treatment areas as compared to control areas.

Table 13 Prevalence of diarrhoea among children 0–35 months

Indicator name	Endline			
	Total	T	C	Diff (T-C)
Children who have suffered diarrhoea in the last two weeks (*children 0–35 months*)	37.8	39.9	35.7	4.2*
N	*5,291*	*2,669*	*2,622*	
Proportion of children (*0–35 months*) who had diarrhoea in the last two weeks and…:				
…suffered from bloody diarrhoea	11.4	12.2	10.7	1.5
N	*1,960*	*998*	*962*	
…sought advice/treatment for diarrhoea from any source	69.7	69.4	70.0	-0.6
N	*1,977*	*1,008*	*969*	
…received ORS at clinic, hospital, health centre	24.0	26.6	21.1	5.5
N	*1,977*	*1,008*	*969*	
…received SSS at clinic, hospital, health centre 1/	5.0	6.4	3.5	2.9*
N	*1,972*	*1,004*	*968*	

Notes: The 'N' shows the number of unweighted observations. Significance levels of the T-C difference are reported with stars: ***Significant at 99.9% level, **Significant at 99% level, *Significant at 95% level.
Data from endline survey only.
1/ SSS

Table 14 displays the proportion of children aged 0–35 months who had an illness in the two weeks preceding the survey. In particular, 40% of children had an illness with fever, while 17% had an illness with a cough. Of those children suffering from either illness, 74% sought treatment/advice from any source, on average.

Table 14 Incidence of different illnesses in children 0–35 months

Indicator name	Endline			
	Total	T	C	Diff (T-C)
In the last two weeks, child (*children 0–35 months*) had:				
Illness with fever	39.7	39.1	40.2	-1.1
N	*5,282*	*2,664*	*2,618*	
Illness with cough	16.5	17.4	15.7	1.7
N	*5,287*	*2,668*	*2,619*	
Sought advice/treatment for fever or cough 1/	74.3	72.9	75.6	-2.7
N	*2,320*	*1,154*	*1,166*	

Notes: The 'N' shows the number of unweighted observations. Significance levels of the T-C difference are reported with stars: ***Significant at 99.9% level, **Significant at 99% level, *Significant at 95% level.
Data from endline survey only.
1/ Of those children who had illness with fever/cough in the last two weeks

3.1.3 Attrition

One concern of panel surveys is the threat posed by attrition bias, i.e. the fact that households drop out of the sample between baseline and endline in such a way that it biases estimates.

Attrition per se should not be a problem if it happens randomly between treatment and control groups, and if it does not occur with such incidence that it affects the representativeness of the remaining sample. Table 3 provided some evidence that attrition was not negligible in our sample (the sample of eligible households decreased by 19% from baseline to endline), thus, not addressing this issue could be problematic for both the point estimates of the descriptive indicators and the estimations of treatment effects presented in Volume 1. Note that Section 3.1.1 presents survey implementation and the reasons why attrition happened in our survey.

To assess whether attrition can introduce bias to our treatment estimates, we run some regressions to test whether it is associated with the treatment assignment. Results presented in Table 15 below show that this is not the case. We run probit regressions in a sample of eligible baseline households on two main outcome variables: the probability of a household being followed up at endline, and the probability of a household being followed up and being eligible for interview at endline (i.e. having a child under three as a household member). For each outcome, we run three specification models without covariates (Specification 1), with state and LGA dummies (Specification 2), and with state, LGA dummies and household-level covariates (Specification 3). All relevant models with some type of covariates show that the treatment assignment (treatment variable) is not significantly correlated with the outcome of interest, suggesting that attrition might not be shifting the composition of our sample differentially across treatment and control groups, once we control for key background characteristics of households (e.g. geographical location (LGA/state)).

Table 15 Attrition regression models

Variables	Household was followed up at endline			Household was followed up and eligible at endline		
	(1)	(2)	(3)	(1)	(2)	(3)
Treatment	-0.331**	0.145	0.208	0.002	-0.222	-0.109
	(0.137)	(0.463)	(0.447)	(0.066)	(0.201)	(0.216)
State /LGA dummies	No	Yes	Yes	No	Yes	Yes
Controls	No	No	Yes	No	No	Yes
Constant	1.898***	1.538***	4.014***	0.851***	0.828***	1.252**
	(0.108)	(0.512)	(0.671)	(0.046)	(0.232)	(0.610)
Observations	*3,355*	*3,215*	*2,978*	*3,355*	*3,355*	*3,211*

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Attrition could still be a problem for our analysis if it affects the representativeness of our sample. For example, if poorer households move away for economic reasons (both in treatment and control areas) and cannot be followed up on, the remaining sample that is being used will be richer than before. Thus, our descriptive and impact estimates would no longer be representative of the original population at baseline.

To account for this problem, we calculate sampling weights corrected for attrition and include these in the calculation of all point estimates of the descriptive indicators presented in Volume 1 and in the robustness checks analysis of the treatment effects presented in Section 3.3 of this report.

To control for attrition in our sample we use inverse probability weights for each household, following an approach outlined in the literature on econometric techniques. These weight households according to their respective probabilities of dropping out of the sample in t+1

(Wooldridge 2002, p. 587 ff.). These weights are calculated by first creating a dummy variable indicating whether a household drops out of the sample in the next time period (t+1). We then run a probit regression with this dummy variable as the outcome variable and use a set of household-characteristics covariates that are thought to be correlated with the probability of dropping out. These regressors were chosen based on theoretical priors as well as using forward and backward stepwise selection models to determine which regressors were most significantly correlated with attrition.

The predicted probability of household attrition in t+1 (endline) obtained from the probit regression becomes the weight for each household at endline. Finally, we calculate the inverse of this weight to obtain our inverse probability weights. Using inverse probability weights ensures that households that have a higher probability of dropping out of the sample are given more weight in the regression in order to compensate for underrepresentation due to attrition. At baseline, each household keeps the original weight not corrected for attrition.

When using attrition correction weights to estimate point estimates of descriptive indicators we lose some additional observations in our sample, given that attrition-corrected weights were constructed controlling for some key observed characteristics, as explained above. Therefore, households that have any missing values in those characteristics are dropped since they do not have an attrition-corrected weight assigned. Table 16 compares point estimates of key outcomes in order to assess whether we introduce additional bias by including attrition-corrected weights due to this loss of observations. We do not observe significance differences in any of the point estimates presented. Thus, our conclusion is that the additional loss of observations caused by using attrition-corrected weights is not problematic.

Table 16 Comparison of prevalences including attrition- and non-attrition-corrected weights

Indicator name	Attrition-corrected weights						Non-attrition-corrected weights					
	Treatment			Control			Treatment			Control		
	Baseline	Endline	Diff (EL–BL)	Baseline	Endline	Diff (EL–BL)	Baseline	Endline	Diff (EL–BL)	Baseline	Endline	Diff (EL–BL)
Exposure to WINNN programme												
Proportion of mothers (*15–49 years)* who ever attended IYCF counselling in the community	7.5	31.5	24.0***	4.3	7.4	3.1**	7.5	30.8	23.3***	4.3	8.0	3.7**
N	*2,833*	*2,303*		*2,833*	*2,235*		*2,833*	*2,395*		*2,833*	*2,370*	
Proportion of children who have ever had their mid-upper arm circumference (MUAC) measured (*children 6–35 months*) 1/	12.9	20.1	7.1***	7.6	7.8	0.3	12.9	20.0	7.1***	7.6	8.3	0.7
N	*2,875*	*2,215*		*2,811*	*2,175*		*2,875*	*2,298*		*2,811*	*2,307*	
Breastfeeding indicators												
Early initiation (< 24 hours) to breastfeeding (children 0–23 months)	64.4	82.8	18.4***	60.2	72.9	12.7***	64.4	82.9	18.5***	60.2	72.6	12.4***
N	*2,190*	*1,784*		*2,212*	*1,754*		*2,190*	*1,845*		*2,212*	*1,871*	
Exclusive breastfeeding (children 0–5 months)	9.2	19.5	10.3***	3.1	7.2	4.1*	9.2	19.3	10.2***	3.1	7.3	4.2*
N	*578*	*453*		*554*	*446*		*578*	*469*		*554*	*479*	
Anthropometric indicators												
Wasted (*6–35 months*)	14.9	17.6	2.7	17.5	17	-0.5	14.9	17.8	2.9*	17.5	17.3	-0.2
N	*2,726*	*2,174*		*2,580*	*2,120*		*2,726*	*2,257*	*2,580*	*2,580*	*2,247*	*2,257*
Stunted (*0–35 months*)	52.1	49.2	-3.0	54.5	52.6	-1.9	52.1	49.3	-2.9	54.5	52.6	-1.9
N	*3,306*	*2,606*		*3,104*	*2,554*		*3,306*	*2,706*		*3,104*	*2,713*	
Underweight (*0–35 months*)	38.7	38.5	-0.2	39.1	37.9	-1.2	38.7	38.9	0.2	39.1	38.3	-0.8
N	*3,329*	*2,641*		*3,201*	*2,571*		*3,329*	*2,740*		*3,201*	*2,732*	

3.1.4 Data collection and data quality assurance

Data collection

Baseline data collection took place in June and July 2013, while endline data collection took place in July and August 2016. Both data collection rounds were completed during the rainy season to avoid bias due to seasonal fluctuations. OPM developed all survey instruments and protocols and oversaw the complete data collection process for this evaluation. At baseline, OPM collaborated with the Nigerian survey company DRMC to conduct the ORIE listing and implement the data collection: DRMC carried out the data collection with supervision from OPM. At endline, OPM led all data collection activities independently.

The listing, baseline and endline surveys were conducted by separate and independent teams. State listing teams were recruited at baseline from experienced personnel native to northern Nigeria, with each team consisting of at least one member from the National Bureau of Statistics. The same listings forms developed at baseline were used at endline to follow up with panel households.

Detailed maps of the EAs were prepared to enable the baseline and follow-up survey teams to identify sampled households. Basic household information to identify the household, as well as the nature of its composition, was collected at baseline and was used at endline to track the same households. In total, 840 EAs were selected at baseline for the study. (See Section 3.1.1 for more detail on sampling.)

The composition of data collection teams at endline was similar to baseline data collection arrangements. Each state team was composed of one state coordinator, one quality assurance officer, and six survey teams (i.e. one for each LGA in a state). Each survey team consisted of: two interviewers (female), one anthropometric specialist (female) and one driver. Interviewers, anthropometric specialists, and quality assurance officers were composed of only female team members who could speak Hausa fluently and who had experience of conducting surveys in northern Nigeria.

Data quality assurance

In addition, several steps were taken by OPM to improve data quality at endline compared to baseline. First, to minimise any inconvenience on the part of the respondent and minimise data-entry and data-processing errors, data collection was conducted at endline through computer-assisted personal interview (CAPI) software.

Second, four levels of data quality assurance were put in place during data collection:

- First, automatic quality check mechanisms were included within the CAPI software. This means that the CAPI software automatically checked things such as the plausibility of values entered, and created error messages for implausible values. For example, if someone entered an age of a respondent that was above 100 years, the CAPI software would mention that this was implausible.
- Second, quality assurance officers were present at all times during fieldwork within the states. Quality assurance officers visited different survey teams and regularly checked the progress of data collection, responded to queries from survey team members, and ensured that fieldwork progressed as planned, in terms of data quality.
- Third, logic and consistency checks of data collected took place on a daily basis. This means that data were uploaded every day onto a server, where datasets were compiled and then fed

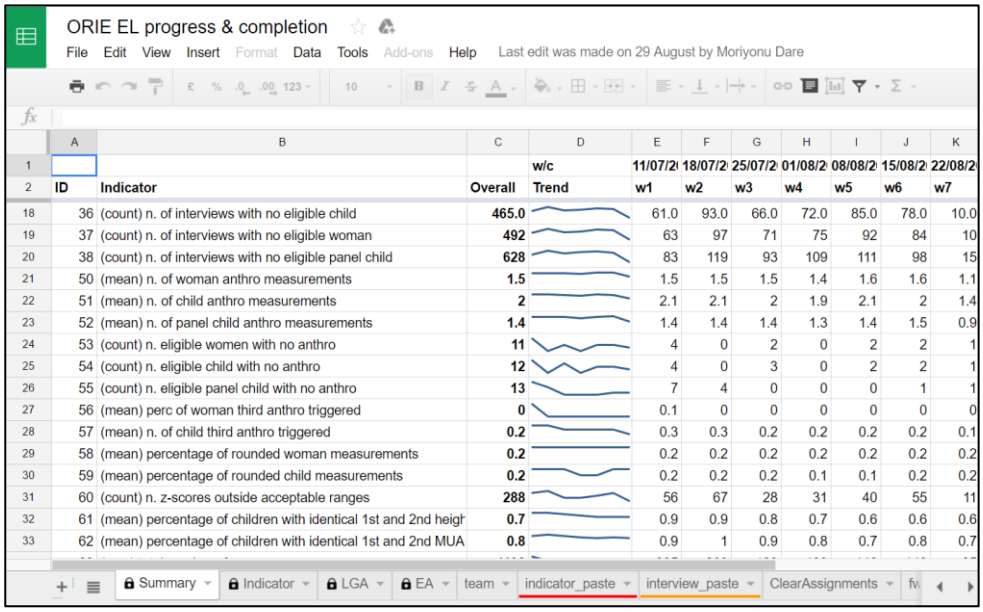
into a programme (Stata do-file) that ran logic and consistency checks on that data. The results of these checks were shared on a daily basis with supervisors and state teams via text messages. Figure 3 is an example of a feedback message shared with the team, where a consistency issue was found with regard to the age at which a mother reported having given birth.

Figure 3 Example feedback message shared with state team

108	MAKE SURE TO PROBE AND CONFIRM THE AGE OF THE PERSON SELECTED AS A HEAD IF THIS SEEMS TOO LOW. ALS
109	**[list interviewers]** had the following inconsistency: The age at which a woman had her FIRST child was below 12. This is very uncommon as most women are not fertile at that age. MAKE SURE TO PROBE AND CONFIRM AGE OF THE WOMAN AND OF HER FIRST CHILD EVERY TIME THE DIFFERENCE
	[list interviewers] had the following inconsistency: The age at which a WOMAN said she gave birth for the first time in the woman questionnaire is inconsistent with the age of her firs

- Fourth, performance checks of the interviewers and the teams took place on a weekly basis by comparing their performance to historical performance and global averages across all the different survey team members. These data were stored and tracked over the entire data collection period using a dashboard developed by OPM. Feedback on performance was shared with the teams on a weekly basis as well. Indicators included in these weekly performance analyses were, for example, mean interview durations, mean answers recorded per minute, mean number of eligible children contacted, percentage of anthropometric measurements (height and weight) rounded to end-digits of 0 and 5, and the percentage of anthropometric measurements of length where the first and second measurement were identical. Significant divergence from global averages or historical performance were then discussed among the team. Figure 4 gives an example of how these indicators were tracked in a dashboard developed by OPM.

Figure 4 Example of continuous tracking of performance indicators for interviewers



	ID	Indicator	Overall	w/c Trend	11/07/2 w1	18/07/2 w2	25/07/2 w3	01/08/2 w4	08/08/2 w5	15/08/2 w6	22/08/2 w7
18	36	(count) n. of interviews with no eligible child	465.0		61.0	93.0	66.0	72.0	85.0	78.0	10.0
19	37	(count) n. of interviews with no eligible woman	492		63	97	71	75	92	84	10
20	38	(count) n. of interviews with no eligible panel child	628		83	119	93	109	111	98	15
21	50	(mean) n. of woman anthro measurements	1.5		1.5	1.5	1.5	1.4	1.6	1.6	1.1
22	51	(mean) n. of child anthro measurements	2		2.1	2.1	2	1.9	2.1	2	1.4
23	52	(mean) n. of panel child anthro measurements	1.4		1.4	1.4	1.4	1.3	1.4	1.5	0.9
24	53	(count) n. eligible women with no anthro	11		4	0	2	0	2	2	1
25	54	(count) n. eligible child with no anthro	12		4	0	3	0	2	2	1
26	55	(count) n. eligible panel child with no anthro	13		7	4	0	0	0	1	1
27	56	(mean) perc of woman third anthro triggered	0		0.1	0	0	0	0	0	0
28	57	(mean) n. of child third anthro triggered	0.2		0.3	0.3	0.2	0.2	0.2	0.2	0.1
29	58	(mean) percentage of rounded woman measurements	0.2		0.2	0.2	0.2	0.2	0.2	0.2	0.2
30	59	(mean) percentage of rounded child measurements	0.2		0.2	0.2	0.2	0.1	0.1	0.2	0.2
31	60	(count) n. z-scores outside acceptable ranges	288		56	67	28	31	40	55	11
32	61	(mean) percentage of children with identical 1st and 2nd heigh	0.7		0.9	0.9	0.8	0.7	0.6	0.6	0.6
33	62	(mean) percentage of children with identical 1st and 2nd MUA	0.8		0.9	1	0.9	0.8	0.7	0.8	0.7

Particular emphasis was put on ensuring the quality of the anthropometric data. At baseline and endline, all anthropometric measurements were made by someone who had previous experience in using such equipment and whose sole responsibility was to take accurate anthropometric measurements. Following standard procedures, specific additional procedures were put in place at baseline and endline to ensure that the anthropometric data were accurately collected and recorded. Children who could not stand on their own underwent a special procedure: their height was measured while lying down (and this was recorded in the CAPI system) and their weight was recorded with the caregiver (one measure of the weight of the child and caregiver together was

taken and then one measure of the weight of the caregiver alone). Additionally, at endline, two measurements of each anthropometric measure were taken (weight, height, MUAC), so as to limit errors. A third measure was randomly triggered to ensure the quality of the data. In order to facilitate age determination at baseline and endline, a seasonal calendar (to find the right time period) and an event calendar (to identify year and, where possible, month) was built to help mothers and interviewers pinpoint the data of birth of the child as accurately as possible. Show-cards were then used to determine the actual age of the child by converting the date into age. Finally, extensive training for enumerators on how to use anthropometric measurement equipment and how to determine age took place, with repeated exercises and tests with both children and mothers.

Training activities

As with the baseline training, the entire endline survey team was brought to Abuja for a training programme that spanned two weeks. ORIE endline quantitative data collection activities were scheduled to begin on 11 July 2016. As part of preparatory activities for this assessment the evaluation team (comprising OPM Oxford and Nigeria staff) trained data collectors, quality assurance officers, and state coordinators, between 22 June and 03 July 2016. Anthropometric specialists were given specialist training through a combination of parallel and joint sessions with the interview and supervisory teams. Anthropometric methods were standardised following methods recommended in the Food and Nutrition Technical Assistance (FANTA) guidelines (Cogill, 2003). All team members then participated in two 'live pilots' that were conducted in Hausa-speaking surrounding areas of Abuja.

The training sessions focused on ensuring participants understood the survey protocol, as well as appropriate methods of administering the different questionnaires – household, mother, child, and community questionnaires. Roles and responsibilities for the fieldwork phase were also clarified and assigned during this process, including the selection and training of some participants to administer the anthropometric tests. Additionally, participants were trained on the features and appropriate use of CAPI devices for data collection.

Instruments

Data were collected using four instruments: one household questionnaire was administered per household. In addition, separate questionnaires were administered for each child aged 0–35 months and each mother of reproductive age (15–49 years) of children 0–35 months in the household. Each child questionnaire was answered by the child's main caregiver, which in the majority of cases was the child's mother. Mother questionnaires were only answered by mothers of the children who were present at the time of the interview, not by caregivers. The child and mother questionnaires included an anthropometric module which was completed by anthropometric specialists. Finally, one community questionnaire was administered for every EA. The full questionnaire can be found in Annex D.2.

Information collected from respondents (main caregivers, mothers or household heads) was based on recall over various time periods. Any recall bias was mitigated through rigorous pre-testing of all survey instruments, the use of standardised methodologies and scales, in-depth enumerator training, and close supervision of the survey work. At endline, a new module was included to capture the exposure to WINNN interventions and some changes were made in the questionnaire as the understanding of the programme and context improved. These new additions or changes in the survey instruments were made after consultation on programme activities with implementing partners, interviews with LGA WINNN focal points and pre-testing of instruments in Katsina.

Moreover, GPS data of all communities and WINNN-supported health facilities (Outpatient Therapeutic Programme (OTP) facilities, stabilisation care facilities and health facilities providing IYCF-related services) in treatment areas were also collected in order to calculate the distance between each community and these health facilities. Distance was defined as geodetic distance: i.e. 'the length of the shortest curve between two points along the surface of a mathematical model of the earth' (Vincenty, 1975).

Note additionally that the IYCF practices and dietary recall sections in the child questionnaire were used to build the key practice indicators in this report, while the anthropometric measurement section was used to build the malnutrition impact indicators. The mother questionnaire and the child questionnaire were also used to build the WINNN exposure indicators. The wealth indicators were built using the wealth assessment section of the household questionnaire. Precise definitions and sources for all key indicators can be found in Annex E.

Finally, it is important to highlight that the determination of the ages of children, in terms of months, can be particularly difficult in this context. As such, an event calendar was developed and age was determined by asking the child's mother and other members of the household to recall major events that occurred around the time of the child's birth. Such events included religious celebrations, the change in season, local elections and significant events such as the death of an Emir or a plane crash. By knowing the date of a number of significant events that occurred in and around the local community, an interviewer was able to triangulate the month and year that a child was born in.

For this survey, an event calendar was produced specifically for northern Nigeria and was tailored to each community by asking respondents of the community questionnaire to inform the survey team of any significant community-level events – such as when the village was flooded. Some households had a vaccination card and even birth certificates, but experience revealed that age determination by event calendar was more accurate as vaccination cards were typically issued to children many months after they were actually born, especially for children not born in a health facility. Birth certificates were even more unreliable as they are typically issued much later due to the administrative and financial costs associated with getting one.

Ethical protocols

Both at baseline and endline, the data used in this analysis were collected following strict ethical standards. Three particular steps taken to ensure that this happened need to be mentioned here:

- First, given the cultural context in northern Nigeria, all interviewers and anthropometric measurement specialists in our survey teams were female. This ensured that interviews with mothers and children could be conducted without respondents feeling uncomfortable about the situation because of the interviewer's gender.
- Second, consent was collected explicitly from all respondents. Please see Annex D.1 for examples of consent forms. Note that those were translated into Hausa and interviewers were trained in explaining the forms to respondents.
- Third, at endline, there was an explicit protocol in place in terms of how to deal with children that were identified to be severely acutely malnourished. First, because data were entered electronically, the CAPI programme automatically identified children who were potentially severely acutely malnourished. This was based on the anthropometric data collected by the anthropometric specialist (i.e. on height/length, weight, age, and the MUAC of each child). Based on this, the CAPI programme automatically notified the interviewer that a child needed to be referred to a primary health care centre for appropriate diagnosis and potential treatment. All survey teams had a list of all primary health centres and community-based management of

acute malnutrition (CMAM) facilities in the area where the survey was being implemented, including a weekly schedule of CMAM days. The team therefore filled in a referral form and handed this over to the caregiver, to give to a health professional at the health centre. The survey team encouraged the caregiver to seek treatment and – where possible – try to provide transport in their own vehicle, which unfortunately was not always possible. For reference purposes, the referral form is reproduced in Annex D.3.

3.1.5 Did mothers experience a recall bias when answering MNCH questions?

At baseline, MNCHW events were in operation in Katsina, Jigawa and Zamfara. The baseline survey was carried out in June/July 2013 and the most recent MNCHW event took place in May 2013. Therefore, we expect any recall bias to be limited given the last MNCHW event took place within two months of the baseline survey.

At endline, MNCHW events were operational in all four states. The endline survey was carried out in July/August 2016 and the most recent MNCHW event took place in June 2016 in Jigawa and in July 2016 in Katsina, Kebbi and Zamfara. As the MNCHW events in Jigawa took place within two months of the endline survey, we do not expect there to be any recall bias for survey respondents in Jigawa. In the other three states, there is a chance that they responded to the endline survey before the July 2016 MNCHW event. For this limited group of respondents – there is a chance of recall bias as the previous round of MNCHW events would have been more than two months ago: in December 2015 in Katsina, in February 2016 in Kebbi and in January 2016 in Zamfara.

To explore this in more details, we created an indicator that captures the number of days that elapsed between the date of the interview of the mother and the last MNCHW event that the mother could have attended. This is a mother's potential bias. We then created a variable that captures whether a mother might be at risk of a potential recall bias (if the interview date was before the MNCHW events took place and therefore the last MNCHW event the mother could have had attended was the November 2015 round) or not (if the mother could have attended the May round of 2016, taking place in June or July). Note that MNCHW events extend over several days and it is impossible to know whether mothers for whom the interview took place during the week of MNCH events attended the event or not before the interview. We considered that all mothers who had an interview during MNCHW could have had attended an MNCHW event during this cycle.

We ran our usual prevalence and impact estimation model on two groups to evaluate the extent of the bias. First, we present the estimates in the full sample as already presented in Table 21 in Section 4.1.3 in Volume I of this report (Quantitative Impact Evaluation of the WINNN Programme – Volume 1, 2017). Second, we present estimates for mothers with no potential recall bias.

Overall, we find that we are not worried about a potential recall bias from the last attendance at MNCHW events. **Error! Reference source not found.** and Table 18 show that estimates on the full sample and on the sample of mothers who could have attended the last MNCHW event (mothers with no potential recall bias) are very similar. Prevalences at endline in both treatment and control areas are within the same range and the significance levels are robust to the exclusion of mothers with potential recall bias. Thus, these findings suggest that, even if we took out mothers with a potential recall bias from the analysis, the results would remain unchanged. Therefore this analysis indicate that there is no evidence of potential recall bias diluting the estimates found in section 3.1.5 in Volume 1 of this report.

Table 17 Recall bias in MNCHW events attendance

Indicator name	Treatment			Control			Impact estimate (S.E.)
	BL	EL	Diff (EL–BL)	BL	EL	Diff (EL–BL)	
Proportion of mothers (*15–49 years*) who have attended the last MNCHW events 1/							
Full sample 2/	5.0	14.8	9.8***	4.8	8.2	3.4*	6.3** (2.0)
N	*2,009*	*1,640*		*2,080*	*1,673*		
No potential recall bias sample (*mothers 15–49 years who could have attended MNCHW events in June/July 2016*)	5.0	17.1	12.0***	4.8	9.1	4.3*	8.1** (2.5)
N	*2,009*	*1,385*		*2,080*	*1,177*		

Notes: The 'N' shows the number of unweighted observations. Significance levels of the T-C difference and DID are reported with stars: ***Significant at 99.9% level, **Significant at 99% level, *Significant at 95% level.
We do not present impact estimates including Kebbi here since Kebbi did not hold MNCH weeks at baseline and therefore data are not comparable from baseline to endline.
1/ At baseline this indicator refers to attendance at last MNCH week. At endline, this indicator refers to attendance at any of the last two MNCH weeks for Katsina and Zamfara, and at the last MNCH week for Jigawa.
2/ Same indicator as in Table 21, Section 4.1.3 in Volume 1 of this report (ORIE Quantitative Impact Evaluation – Volume 1 of the final report, 2017).

Table 18 Recall bias in receiving Vitamin A drop at MNCHW events

Indicator name	Treatment			Control		
	BL	EL	Diff (EL–BL)	BL	EL	Diff (EL–BL)
Proportion of mothers (*15–49 years who went to the last MNCHW events*) whose children received Vitamin A drop 1/						
Full sample 2/	93.0	76.0	-17.0***	92.7	78.5	-14.2*
N	*98*	*267*		*88*	*136*	
No potential recall bias sample (*mothers 15–49 years who could have attended MNCHW events in June/July 2016*)	93.0	78.4	-14.6**	92.7	85.8	-6.9
N	*98*	*246*		*88*	*105*	

Notes: The 'N' shows the number of unweighted observations. Significance levels of the T-C difference and DID are reported with stars: ***Significant at 99.9% level, **Significant at 99% level, *Significant at 95% level.
We do not present estimates including Kebbi here since Kebbi did not hold MNCH weeks at baseline and therefore data are not comparable from baseline to endline. We do not present impact estimates due to the small size of the sample.
1/ At baseline this indicator refers to attendance at the last MNCH week. At endline, this indicator refers to attendance at any of the last two MNCH weeks for Katsina and Zamfara, and at the last MNCH week for Jigawa.
2/ Same indicator as in Table 24 in Section 4.1.3 in Volume 1 of this report (ORIE Quantitative Impact Evaluation – Volume 1 of the final report, 2017).

Similarly, in

Table 19, prevalences at endline in both treatment and control areas are within the same range and, the significance levels are robust to the exclusion of mothers with a potential recall bias. Therefore, we conclude that we are not worried about a potential recall bias arising from the last attendance at an MNCHW event.

Table 19 Recall bias in regard to receiving other services at MNCHW events (endline only)

Indicator name	Endline			
	Total	Treatment	Control	Diff (T-C)
Proportion of mothers (*15–49 years, who went to the last MNCHW events*) who received:				
Tetanus toxoid vaccine				
Full sample 1/	39.8	38.7	41.7	-3.0
N	*441*	*292*	*149*	
No potential recall bias sample (*mothers 15–49 years who could have attended MNCHW events in June/July 2016*)	38.6	35.3	44.8	-9.5
N	*389*	*271*	*118*	
Long-lasting insecticidal nets				
Full sample	51.2	52.8	48.5	4.3
N	*450*	*299*	*151*	
No potential recall bias sample (*mothers 15–49 years who could have attended MNCHW events in June/July 2016*)	47.6	49.6	44.0	5.6
N	*398*	*278*	*120*	
Proportion of mothers (*15–49 years, who went to the last MNCHW events*) whose children received the following:				
Deworming pills				
Full sample 1/	61.4	61.1	62.0	-0.9
N	*438*	*292*	*146*	
No potential recall bias sample (*mothers 15–49 years who could have attended MNCHW events in June/July 2016*)	63.9	62.8	65.9	-3.1
N	*386*	*271*	*115*	
Undernutrition examination with MUAC				
Full sample 1/	42.1	45.0	36.9	8.1
N	*444*	*296*	*148*	
No potential recall bias sample (*mothers 15–49 years who could have attended MNCHW events in June/July 2016*)	41.2	42.7	38.5	4.2
N	*392*	*275*	*117*	

Notes: The 'N' shows the number of unweighted observations. Significance levels of the T-C difference and DID are reported with stars: ***Significant at 99.9% level, **Significant at 99% level, *Significant at 95% level.
Data from endline survey only since this information was not known at baseline.
1/ Same indicator as in Table 25 in Section 4.1.3 in Volume 1 of this report (Quantitative Impact Evaluation of the WINNN Programme – Volume 1: Operations Research and Impact Evaluation, 2017).

3.1.6 Questionnaires and indicators definitions

From baseline to endline the questionnaire was adapted to facilitate answers from respondents based on experience from baseline, as well as to collect new information that could help us to understand the broader context in which the WINNN programme was implemented. As the WINNN protocol evolved over time, the questionnaire also evolved to reflect those changes. The final questionnaire used for the impact evaluation is presented in Annex D.2.

The endline questionnaire in Annex D.2 is the main data source for the indicators presented both in Volume 1 and Volume 2. A list of key indicators and their definitions is presented in Annex E.

3.2 Anthropometric data quality assessment

3.2.1 Introduction

This section presents the results of a data quality assessment conducted to understand the quality of the anthropometric data from the ORIE baseline and endline surveys. The analysis was

performed by using information on age, sex, height and weight for children aged 0–35 months in the ORIE survey areas. Following the approach used in the ORIE baseline analysis, quality criteria analysed in this assessment include age and sex ratios, standard deviations of anthropometric z-scores, digit preferences for weight and height variables, clumping of age variables, and the distribution of z-scores (including the presence of outliers) (ORIE Baseline, 2013).

Where possible, and where it is useful, we compared results from the ORIE sample with a quality assessment of the data from the National Nutrition and Health Survey (NNHS) of 2015, in order to benchmark our results against a nationally representative survey. The NNHS was carried out by the National Bureau of Statistics: 20,060 children from 37 states in Nigeria were interviewed. This provides a useful comparator for the ORIE sample.

3.2.2 Balancing of key demographic indicators

Table 20 presents the estimated age and sex ratios for children aged 0–35 months in our data and NNHS (2015), along with proposed Emergency Nutrition Assessment (ENA) guidelines (UN High Commissioner for Refugees (UNHCR), 2013). The age ratio is defined as the number of children aged 0–17 months over the number of children aged 18–35 months. It gives an indication of whether our age distribution is skewed towards older or younger children. The sex ratio is defined as the number of male children over the number of female children. For both, if the sample is perfectly balanced, the value should be 1.0.

Sex and age ratios for both ORIE surveys (baseline and endline) lie within the ENA recommended intervals. The sex ratio for ORIE stood at 1.02 and 1.04 for baseline and endline, respectively, while the age ratio for ORIE stood at 1.04 and 1.14 for baseline and endline, respectively. Overall, this rough analysis implies that the ORIE data are relatively balanced in terms of sex of children and the proportion of younger versus older children.

Table 20 Age and sex ratios

	Sex Ratio	Age Ratio
Proposed ranges by ENA	[0.80-1.20]	[0.78-1.18]
ORIE baseline sample	1.02	1.04
ORIE endline sample	1.04	1.14

Notes: The sex ratio **is defined as the number of male children over the number of female children. For ORIE data, the** age ratio **is defined as the number of children aged 0–17 months over the number of children aged 18–35 months.**

3.2.3 Analysis of the age variable

In order to assess whether our age data for children (in months) show signs of preference for certain ages, Figure 5 presents the distribution of children's ages in months in the ORIE sample for both baseline and endline data. Without preferences for certain ages, we would expect a uniform distribution of children along the full length of the distribution, i.e. from 0 to 35 months. However, panel a) demonstrates that the ORIE baseline data show signs of age-clumping at around 12 and 24 months. This means that there is a preference for children to be reported as being exactly one and two years old. Such age-clumping is also present in the endline data, but to a lesser degree.

Figure 5 Analysis of the age variable: months

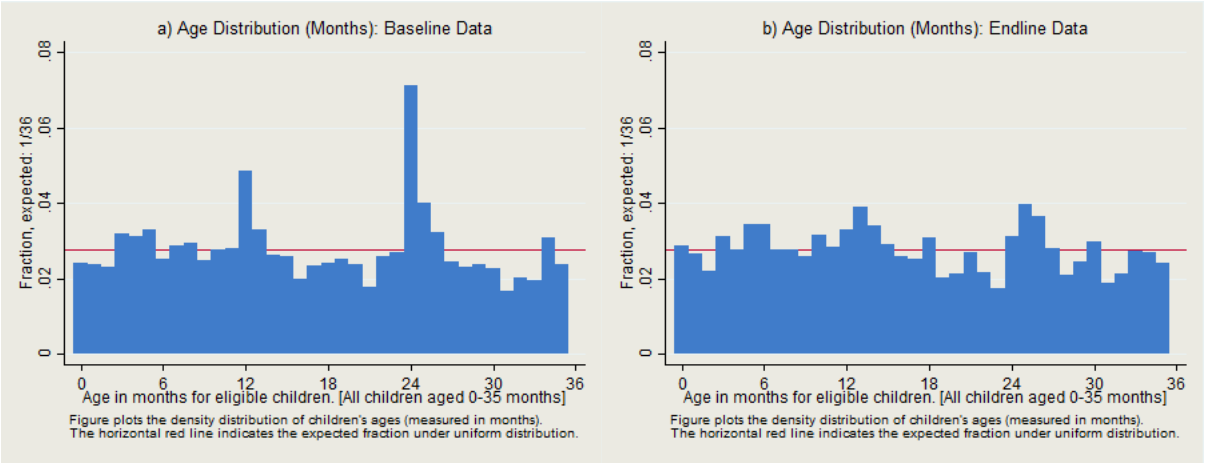
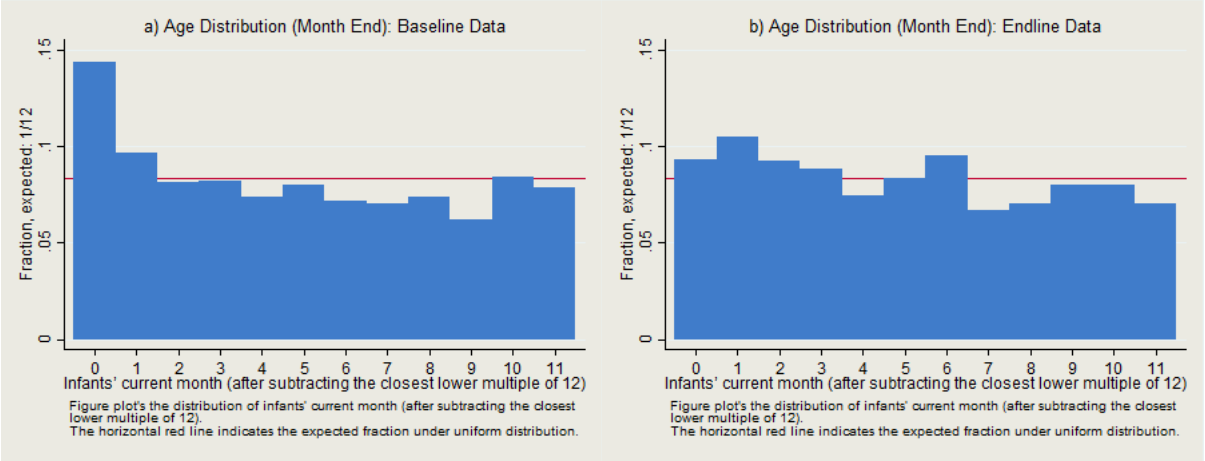


Figure 6 shows the distribution of children's ages across the number of months that are left when the closest lower multiple of 12 is subtracted from children's ages in months. That is, the figure shows the distribution of children's ages in months when full years have been subtracted. This gives an indication for whether there is a preference for children's ages to be reported in yearly patterns. The first panel again shows that in the baseline data there is a clear preference for children's ages to be reported in full years (digit 0). This preference is less pronounced in the endline data, where the bars are more equally distributed, indicating an improvement in the way that children's ages were reported between baseline and endline surveys.

Figure 6 Analysis of the age variable: month end



To further corroborate our findings statistically, we run a Pearson's chi-square goodness of fit test to the distribution of the variable plotted in Figure 6. This test allows us to check whether the observed distribution of children across its values is significantly different from a theoretically expected uniform distribution.

Using the Stata routine 'mgof' and controlling for interviewer cluster levels, we obtain a corresponding test statistic of χ^2 = 405.1 for the baseline data and χ^2 = 102.1 for the endline data, with a p-value of under 0.001 for both datasets (Table 21, row 4). This means that in both cases we reject the null hypothesis of observed frequencies corresponding to an expected uniform distribution, which confirms the results of our visual analysis that age-clumping is present both in the endline and baseline data. Note, however, that the test statistic at endline is lower than that for baseline data. This implies that we reject the null hypothesis of a uniform distribution for endline with less confidence than baseline, again implying better data quality in the endline data.

Table 21 Age-clumping

Age measure	Baseline	Endline
N	6,706	5,540
Test statistic	405.1	102.1
F-value	16.2	7.5
P-value	0.0	0.0

Notes: This table presents results of a Pearson chi-square goodness of fit test applied to the 'months left' variable plotted in Figure 6, panel b), which allows us to test whether the observed distribution differs statistically from the expected uniform distribution.

Collecting age data in months from children in the field is a very difficult exercise. For the survey, age data were collected from vaccination cards and other 'hard' sources and interviewers were extensively trained in the use of event calendars to help caregivers indicate the age of their child when such sources were not available. The analysis above shows that there was an improvement in data collection between the baseline and endline surveys, when looking at the age distribution of our sample, but that age-clumping persists in both cases. This is a common problem in nutrition surveys, including in recent surveys implemented in Nigeria (Kosternians, 1952; NNHS 2015, p. 34).

3.2.4 Analysis of height and weight variables

3.2.4.1 End-digit preference in height and weight variables

This section focuses on the quality of the anthropometric measures of weight and height of children in the ORIE survey. These characteristics are continuous measures of the human body that are determined by a complex interplay of genetic predispositions and the environmental context (Silventoinen, 2003). Scholars generally agree that the distribution of height and weight in a population follow a Gaussian or log-normal distribution (Tanner, 1981; Snedecor and Cochrane, 1989; Limpert *et al*., 2001).

Anthropometric measures collected in the ORIE baseline and endline surveys used similar equipment and were measured to the nearest 0.1 kg for weight and 0.1 cm for height. Note that standard procedures are such that an infant's weight was measured together with the caregiver in cases where the infant was not able to stand up alone using the tare function of the scale. Similarly, height for children aged less than 24 months was measured while lying down ('length') using the measuring board, while older children's height was measured standing up using the mobile stadiometer. Using this precision of one decimal digit in the measure of anthropometrics, we would expect a uniform distribution of these end-digits in the weight and height measures.

Figure 7 and Figure 8 illustrate the distribution of the end-digit of children's weight and height or length measurements across both baseline and endline data. The expected uniform distribution is illustrated through the red horizontal line in the histograms. In addition, the plot provides information about the number of measures taken in the respective sample (N), the test statistic of a Pearson Chi-square test with the null hypothesis of a uniform distribution across digits, as well as the corresponding p-value adjusted for enumerator/anthropometric specialist clusters.

Figure 7 End-digit distribution of decimals in weight variable

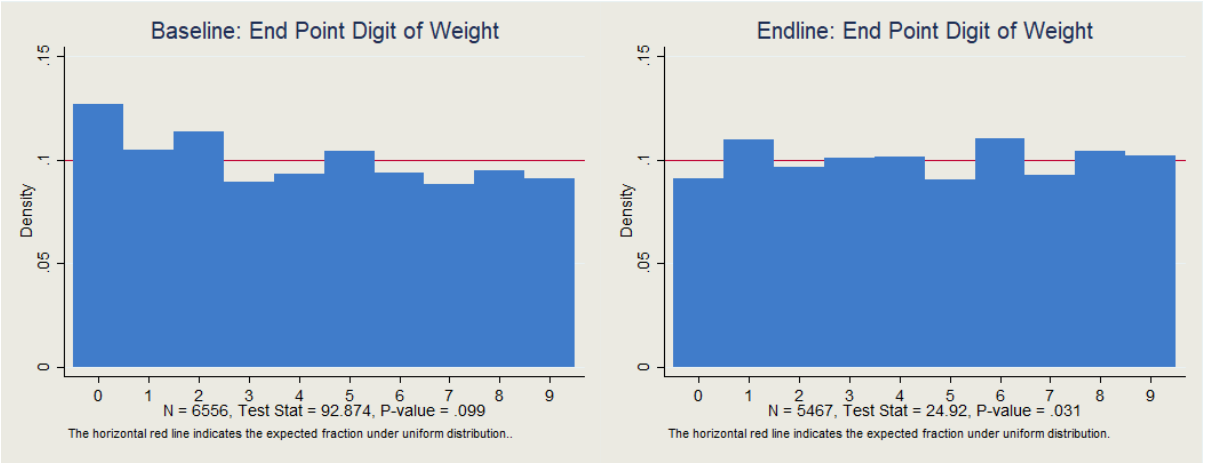
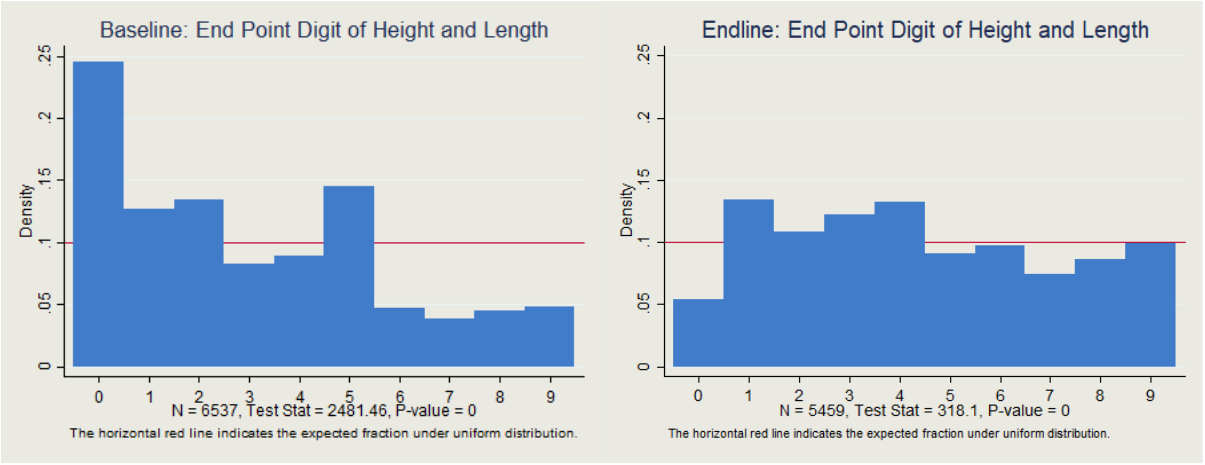


Figure 8: End-digit distribution of decimals in height variable



A visible inspection shows that end-digit preference is present for both the weight and height variables in the baseline data, although it seems to be more pronounced in the case of the latter. For example, at baseline, measures with end-digits .0 and .5 for height and length measures are much more prevalent than would be expected under a uniform distribution, indicating preference for such measures. A visible inspection of the graphs for endline data suggests that this issue of end-digit preference is less prevalent here. Again, this indicates an improvement in data quality from baseline to endline.

For the weight variable (see Figure 7) we obtain a test statistic of χ^2 = 92.87 (p=0.099) for baseline data, suggesting that we fail to reject the null hypothesis of a uniform distribution for the weight variable at baseline when applying a threshold of p=0.05. For endline data on weight, we obtain a test statistic of χ^2 = 24.92 (p=0.031) and thus reject the null hypothesis of a uniform distribution of end-digits when adjusting for anthropometric specialist clusters and applying the same threshold of p=0.05. This implies that end-digits of endline data on weight are slightly less uniformly distributed.

For the height variable (see Figure 8), we obtain a test statistic of χ^2 = 2481.46 (p<0.001) for baseline and χ^2 = 318.1 (p<0.001) for endline, therefore providing stark evidence against the null hypothesis of a uniform end-digit distribution in both cases. This analysis confirms the visual analysis that end-digits of weight data are more uniformly distributed in our samples compared to height data, and that endline data on length and height of children are of slightly better quality than the baseline data.

This analysis therefore suggests that in particular the measurement of height was difficult in the ORIE survey, and that this issue is more severe in the baseline data. Additional analysis (not presented here) indicates that this problem of end-digit preferences exists irrespective of a child's age. Possible explanations for the comparatively high prevalence of end-digit preference in height are mainly related to the measurement process: whereas weight measurement is taken using an electronic scale from which digital measures can easily be read, height measurement is taken using length and height boards. Children need to be put in the right position and marks on the measuring boards might be difficult to read, hence leading to implicit rounding.

This finding could imply that nutrition status classifications of children relying on height measures could be affected by measurement error. These include height-for-age and weight-for-height, or stunting and wasting measures. In particular, assuming that end-digit preference is random or a symmetric rounding error, the prevalence of malnutrition based on these measures will be overestimated due to larger tails of the resulting z-score distribution (Kostermans, 1994). We therefore investigate whether this could be problematic in the following sections.

3.2.4.2 Influence of end-digit preference on estimated prevalence of malnutrition

As a consequence of the detected age-clumping and end-digit preference, estimating nutritional indicators based on the combination of age and anthropometric information might produce upward biased prevalence measures of malnutrition. In order to shed light on this overestimation hypothesis the next set of figures displays malnutrition estimates of children with and without end-digits of 0 and 5 across several disaggregating characteristics. If estimates of malnutrition prevalence are roughly comparable, we conclude that, overall, our estimates will not be biased.

Table 22 presents the prevalence of malnutrition by end-digit preference for underweight and stunting for both baseline and endline surveys. A two-sample t-test of significant difference between prevalence estimates with and without end-digit preference was conducted (p-values reported in Table 22). We found p-values for these tests at baseline to be less than 0.05, i.e. indicating that the difference between the two groups was significantly different from zero. Thus, we reject the null hypothesis of no significant difference. The results show, as expected, that prevalences are slightly higher among the group of children with end-digits of .0/.5 in the underlying anthropometric measurements. This does not hold for endline values, where end-digit preference or age-clumping was less prevalent. Note that results presented further below (Figure 11) show that this does not present a risk to our impact evaluations strategy.

Table 22 Malnutrition prevalence by end-digit preference

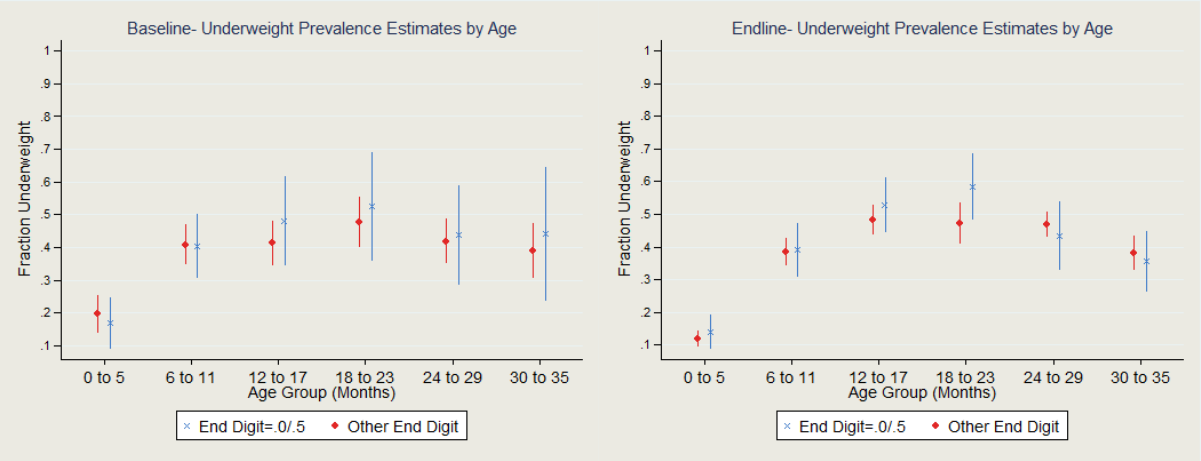
	Baseline			Endline		
	End-digit .0/.5	Other end-digits	P-value	End-digit .0/.5	Other end-digits	P-value
Underweight	44.3	38.9	0.0002	39.4	39.1	0.8589
N	2,497	3,913		782	4,637	
Stunted	55.8	52.8	0.0192	49.5	51.0	0.4466
N	1,510	5,020		993	4,479	

Notes: Prevalence estimates in columns 1, 2, 4 and 5 are in percentages (%).

Figure 9 and Figure 10 display the fraction of children underweight and stunted, respectively, across age groups, comparing children with end-digits 0 and 5 in their height or weight measure, respectively, to children with other end-digits in this measure. The figures include 95% confidence intervals that allow us to assess whether the fractions across end-digit preference and non-end-digit preference are similar in a statistical sense or not.

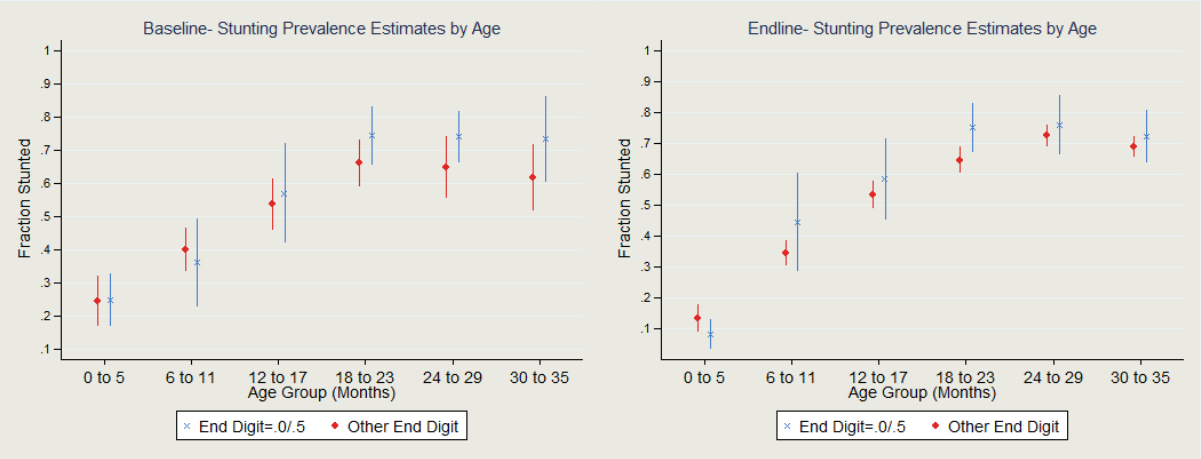
Figure 9 shows that the share of children considered to be underweight within a specific age group differs slightly across end-digit preference groups for weight. Note, however, that the confidence bounds overlap across age groups for both baseline and endline data. This means that prevalence estimates do not differ significantly across the two groups, irrespective of the age group looked at and for both baseline and endline data. This finding, in turn, leads us to conclude that end-digit preference does not seem to be problematic for malnutrition estimates based on weight measures.

Figure 9 Influence of end-digit preference on estimated prevalence of underweight



Similarly, Figure 10 compares the prevalence of stunting for end-digit and non-end-digit preference in height sub-populations across the different age groups. In general, again, confidence intervals overlap when comparing estimates across the two groups at any age group, which indicates that end-digit preference does not seem to be biasing malnutrition prevalence estimates significantly, although differences between the two groups are larger than for the underweight estimates presented above. This is in line with evidence presented in the previous section that end-digit preference is more problematic in height/length measures than in weight measures.

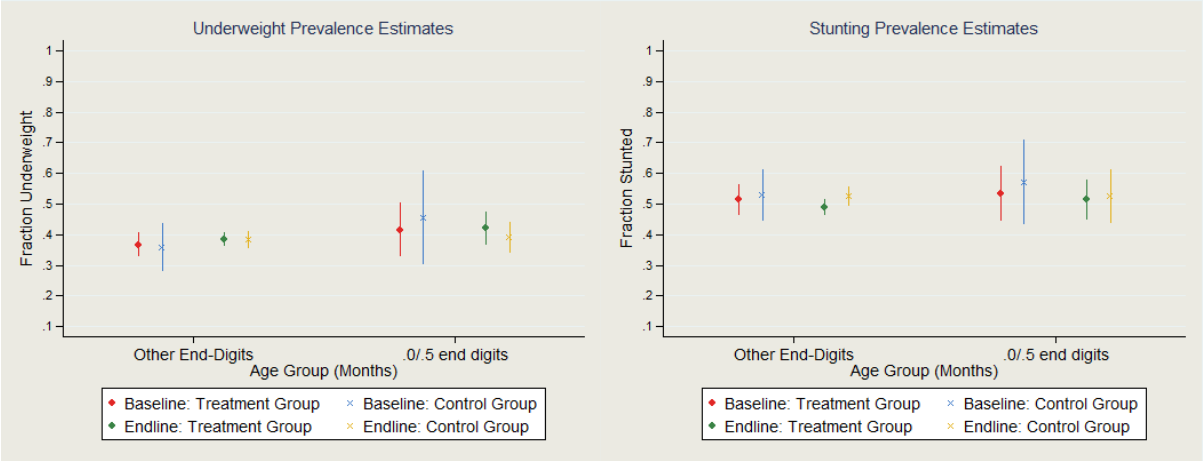
Figure 10 Influence of end-digit preference on estimated prevalence of stunting



As an additional data quality check, and to explore the potential influence of end-digit preference on a DID impact estimate, we compare malnutrition prevalence estimates for observations with end-digits 0 and 5, and without such end-digits, across treatment and control areas and baseline and endline data in Figure 11. If there was a significant differential effect of end-digit preference on malnutrition estimates across treatment and control groups, and across time, this would imply that our impact estimates would potentially be biased due to this preference.

The analysis reveals that, as expected, point estimates for observations with end-digits 0 and 5 are generally higher than estimates without such end-digits. At the same time, confidence intervals comfortably overlap across treatment and control, and both at endline and baseline. We conclude from this that there is little evidence of potential biases in our impact estimates due to end-digit preference.

Figure 11 Influence of end-digit preference on prevalence of malnutrition by treatment and control groups



3.2.5 Analysis of z-scores and undernutrition prevalence rates

The malnutrition indicators investigated graphically in the last section, and presented in Volume 1 of this report, are based on the growth standards published by the WHO in 2006 (WHO, 2006). Specifically, the indicators are derived from a comparison of a child's age and anthropometric information with a global reference population capturing optimal child growth.[2] The present study uses three nutritional indicators: weight-for-age, height-for-age and weight-for-height. Each of these indicators is expressed in standard deviations unit differences from the median of the representative comparison distribution and thus represents a standardised measure or 'z-score'. The international literature on measurement of anthropometric nutritional status suggests that where measurement is accurate, such z-scores exhibit certain distributional features. Hence, looking at the actual distribution of z-scores in our sample can yield insights into the quality of the ORIE data.

3.2.5.1 Standard deviation of estimated z-scores

Column 1 of

[2] Similarly, Figure 9 compares the prevalence of stunting for end-digit and non-end-digit preference sub-populations across the different age groups. There seems to be an overestimation of stunting prevalence for older age groups in the ORIE baseline data in the case of end-digit preference for children above the age of 18 months.

Table 23 presents the acceptable ranges proposed by the WHO (1995) for the standard deviation of each one of the z-scores of interest. Column 2 of

Table 23 presents the updated ranges calculated by Mei and Grummer-Strawn (2007) using the more recent 2006 WHO growth standards and by analysing a variety of existing anthropometric datasets. These are ranges within which, empirically, z-score standard deviations are normally expected to fall. Column 3 and 4 present results for ORIE baseline and endline data, respectively. Similarly, column 5 presents the standard deviation of z-scores used in the NNHS.

Table 23 Standard deviations of z-scores

	WHO (1995)	Mei and Grummer-Strawn (2007)	ORIE Baseline	ORIE Endline	NNHS 2015
Weight-for-age	[1.00-1.20]	[1.17-1.46]	1.56	1.41	1.11
Height-for-age	[1.10-1.30]	[1.35-1.95]	1.80	1.66	1.31
Weight-for-height	[0.85-1.10]	[1.08-1.50]	1.52	1.31	1.05

Notes: This table presents the acceptable ranges proposed by the WHO (1995) and Mei and Grummer-Strawn (2007) for the standard deviation of the z-scores for weight-for-age, height-for-age, and weight-for-height.

Comparing the standard deviations in columns 3 and 4 of

Table 23 (i.e. of ORIE data), with the proposed ranges in column 1 and 2, we find that standard deviations of anthropometric z-scores for both ORIE baseline and endline data are higher than acceptable based on WHO (1995), irrespective of the specific z-score looked at.

However, when comparing the results from ORIE data with ranges suggested by Mei and Grummer-Strawn (2007) we find that our standard deviations are larger than the upper end of these ranges for weight-for-age and weight-for-height at baseline only. The standard deviations of endline z-scores fall comfortably within the suggested ranges. Similarly, the standard deviations of z-scores for NNHS data lie within WHO (1995) and Mei and Grummer-Strawn (2007) recommended ranges, but are smaller than for the ORIE data.

Because measurement error is positively related to the standard deviation of z-scores, i.e. higher measurement error leads to higher standard deviations, this analysis suggests that, first, data quality in ORIE improved between baseline and endline, and, second, that at baseline measurement error was relatively high compared to what can be found in the literature. Note, however, that other surveys implemented in northern Nigeria exhibit similar levels of z-score standard deviations, in particular when related to height measurements. For example, in the NNHS data, the standard deviation values of z-scores for height-for-age are 2.22 in Katsina, 2.21 in Kebbi, 2.21 in Zamfara, and 2.4 in Jigawa (NNHS 2015).

3.2.5.2 Presence of outliers in estimated z-scores

Having investigated malnutrition prevalence estimates and the effects of data quality on these, this section considers the presence of outliers that are excluded from the above prevalence estimations due to being outside of the acceptable ranges according to WHO standards. These acceptable ranges are defined as follows:

- [-6, 5] for weight-for-age (underweight) z-scores
- [-6, 6] for height-for-age (stunting) z-scores
- [-5, 5] for weight-for-height (wasting) z-scores.

Z-scores that lie outside of these ranges are assumed to be outliers that result from extreme measurement errors that make z-scores implausible.

Table 24 presents the percentage of z-scores lying outside of these WHO-acceptable values, along with the acceptable fraction of outliers for each malnutrition indicator across the different samples. Specifically, the table reports the proportion of z-scores lying above and below the respective cut-offs. As a general note for symmetric ranges, a higher proportion of outliers lying under the respective threshold as compared to the share lying above the acceptable threshold suggests a z-score distribution that is skewed to the left.

Table 24 Percentage of z-scores lying outside of the acceptable ranges

	Weight-for-age (%)				Height-for-age				Weight-for-height			
	> 5	< -6	Total	**WHO guidelines**	> 6	< -6	Total	**WHO guidelines**	> 5	< -5	Total	**WHO guidelines**
Baseline	0.97	0.95	1.92	**1.00**	0.42	3.31	3.73	**5.00**	2.15	1.49	3.64	**3.00**
Endline	0.27	0.34	0.61	**1.00**	0.16	1.35	1.51	**5.00**	0.49	0.65	1.13	**3.00**

Notes: This table presents the percentage of z-score values lying outside of the acceptable ranges established by the WHO, that is: [-6, 5] for weight-for-age (underweight) z-scores
[-6, 6] for height-for-age (stunting) z-scores
[-5, 5] for weight-for-height (wasting) z-scores

Overall, the results presented in Table 24 show that the proportion of total outliers present in the ORIE data exceeds WHO guidance in two cases only: weight-for-age and weight-for-height at

baseline. As in the previous analysis, this suggests that measurement error is more present in these z-scores and that data quality improved between baseline and endline.

3.2.6 Conclusions

To summarise, the following key findings emerge from this data quality analysis:

- The quality of anthropometric data, as measured by the presence of measurement errors, age-clumping, end-digit preference, and distributional analysis of z-scores, improved between the ORIE baseline and endline surveys.
- We find little evidence for severe data quality issues, generally. However, we do find evidence for issues of age-clumping among children, i.e. the preference of reporting full or rounded years as children's ages, and of end-digit preference in height measurements. Both are relatively more common in the baseline data.
- We do not find that this has differential effects on malnutrition prevalence estimates across treatment and control groups, and baseline and endline data. We therefore do not find evidence for concerns regarding our impact evaluation strategy.
- Finally, when looking at z-score distributions and the prevalence of implausible outliers, we similarly find evidence for relatively higher levels of measurement error in baseline data, and hence an improvement between baseline and endline data.

3.3 Impact evaluation strategy

3.3.1 DID estimation in the context of the present evaluation

As mentioned in Volume 1, given the nature of the WINNN-supported interventions (the allocation of treatment and control LGAs not being random and various components of the programme being implemented at different stages), a simple comparison of treatment and control areas at endline is not a viable strategy to robustly estimate the effect of the WINNN-supported interventions. There are likely to be systematic differences between treatment and control areas, other than the treatment assignment, that would bias the impact estimates.

We therefore use a DID approach as our main impact identification strategy, which is a suitable approach to deal with differences in unobservable characteristics across treatment and control groups that follow the same trend over time. This approach measures the differential outcome between estimates in treatment and control areas over time (double-difference). The underlying assumption for this approach to work is that the difference in outcomes between treatment and control areas would have been constant over time had it not been for the WINNN-supported interventions. Through this study design, the control group is an acceptable counterfactual provided that any differentials in changes over time are only due to the implementation of the WINNN-supported programme and not due to any other factors, such as, for example, the differential scale-up or roll-out of nutrition-related programmes in or around the evaluation areas.

Please note that Volume 1 provides some detail on our DID estimation strategy. Hence, this section only provides additional technical information, and focuses on presenting additional robustness check results.

Impact estimates and regression specification

General regression specification

As discussed in Volume 1, treatment status in this evaluation was assigned geographically to treatment and control LGAs. This means that anyone living in treatment areas is considered as 'treated', regardless of whether that individual has actually had an interaction with components of the interventions supported by WINNN. Therefore, following from our approach defined in the inception phase of this evaluation, estimates presented in this report are estimates of the ITT effect: this is the average effect on the relevant population in treatment areas, irrespective of whether individuals have been in touch with the WINNN programme.

The ITT estimation is implemented using a regression framework that takes into account the sampling structure of the data, weights and household fixed effects, and includes covariates to control for confounding factors.

The main regression specification is estimated as shown in equation (1):

$$(1) \qquad y_{iht} = \alpha_0 + \alpha_1 Post_t + \alpha_2 Treat_h + \alpha_3 Treat_h * Post_t + \beta X_{iht} + c_h + u_{iht}$$

where y_{iht} is the outcome variable for individual i (e.g. children aged 0–35 months) of household h in time t. $Post_t$ is a dummy to control for a time effect taking the value of 0 for baseline and 1 for endline. $Treat_h$ is a dummy that captures the allocation of treatment and takes the value 1 for households in the treated LGAs and 0 for the households in the control LGAs. Note that in the context of this evaluation, treatment status does not vary over time since treatment is fixed – households have been assigned to either treatment or control groups from the beginning of this evaluation. X_{iht} are individual, household and village-level covariates that vary over time and are expected to be related to programme outcomes: c_h represents household-level time-invariant characteristics (observable and unobservable); and u_{iht} is a random error term. The effect of the WINNN intervention on the outcome y_{iht} is given by the α_3 coefficient, i.e. the coefficient on the interaction of the treatment status with the time dummy. Note that this is the standard DID specification augmented by household fixed effects (Wooldridge, 2012, p. 455).

It is important here to highlight that equation (1) can be specified with a general error term v_{ith}-, where $v_{ith} = c_h + u_{ith}$.

$$(2) \quad y_{iht} = \alpha_0 + \alpha_1 Post_t + \alpha_2 Treat + \alpha_3 Treat * Post_{ht} + \beta X_{iht} + v_{iht}$$

For the estimation of α_3 in equation (2) to be unbiased, the error term v_{ith} needs to be uncorrelated with the rest of the regressors. However, in the context of this evaluation we know there is a correlation between c_h and the covariates (such as unobservable household characteristics). Therefore, we avoid this particular bias by controlling for household varying but time-constant characteristics using a household-level fixed effects estimation, thereby controlling for c_h, i.e. time-constant household-level characteristics. By doing this, we are left with an error term u_{ith} that is assumed to be uncorrelated with the rest of the regressors. For identification purposes, we therefore employ the standard strict exogeneity assumption, which in the present case implies that a household-level fixed effects estimation that additionally includes appropriate covariates allows us to control for all relevant confounding factors (Wooldridge, 2012, p. 467). Note that we deal with covariate selection in more detail below.

Clustering of standard errors

An additional issue in this context is estimating standard errors the correct way. In general, we implemented fixed effects regressions with clustered standard errors at the EA level. However, estimating standard errors in this impact evaluation is not straightforward due to the complex structure of the variance-covariance matrix of the error term in (1). There are several issues that arise when thinking about estimating standard errors in this context:

1. First, the survey design includes a complex sampling structure as described in Section 3.1.1. The PSUs were EAs (community level). The second sampling stage randomly selected household within each EA which had a child under the age of three. Hence, clustering within PSUs needs to be taken into account. Note that clustering at this level does not take into account higher-level clustering, i.e. LGA-level clustering.

2. Second, the impact evaluation literature recommends clustering standard errors at the level at which the treatment is assigned, in order to account for correlation among units within treatment assignment groups. Here, treatment was allocated at the LGA level, which would imply that standard errors could be clustered at this level. However, the literature also recommends that a higher level of clustering can be imprecise when the sample has a small number of clusters (such in our case, where we have 24 LGA clusters) (Cameron *et al*., 2008).

3. Third, this evaluation strategy is based on a panel of households, which imposes additional assumptions on the error term. The panel structure of our data implies an inter-temporal correlation for observations from the same households. Hence, the standard approach in a panel analysis would be to cluster at the level of the panel observations. In our case, this would imply clustering at the household level. Note that clustering at this level does not take into account higher-level clustering, i.e. either EA or LGA level.

As said above, for the purposes of this impact evaluation and the main results presented in Volume 1, and in order to take into account the complex sampling strategy implemented, we cluster standard errors at the PSU level, i.e. the EA level, in all main specifications. Note that this, in effect, also takes into account lower-level correlations between observations, such as, for example, inter-temporal correlations within households. It does not take into account higher-level clustering, i.e. LGA-level clustering mentioned in point 2 above. In our specifications, we decided to not cluster at the LGA level, given that (i) the whole evaluation was planned and the sampling design implemented assuming EA-level clustering, and (ii) evidence for these estimations is imprecise when there is a small number of clusters. However, in Section 3.3.2 below we explore how changing this to household- and LGA-level clustering changes our estimates.

Choosing the covariates

For each key impact indicator and in order to check the consistency and robustness of our findings to changes in the estimation procedures, we ran estimations using different regression specifications. These are described in more detail in the next subsection. Results presented in Volume 1 correspond to our preferred fixed effects estimation procedure.

Selection of covariates to be included in the estimation models was done applying theoretical considerations and employing a selection algorithm known as stepwise regressions. This approach allows us to identify variables that vary significantly between treatment and control groups, and which are also correlated with the outcome of interest. Note that in implementing this approach we follow the double selection literature (Belloni *et al.,* 2014).

There are two stepwise regression approaches that can be employed for this: backward and forward stepwise regression. The underlying idea behind both approaches is to check each

covariate, step-by-step, for significant correlation with the outcome and treatment assignment variables. We are looking for such a correlation because variables that possibly bias our impact estimates (i.e. that are confounders) will have some relation to both the treatment status and the outcome we are looking at.

Backward selection starts with the full set of covariates, i.e. a regression including all variables, and then discards the term that is least significantly correlated with the dependent variable. It continues to do so until all variables that are uncorrelated with the dependent variable are discarded. Forward selection, instead, starts with an empty set of covariates, i.e. a regression on a constant, and then checks the significance of each covariate if it is included in the regression. It then adds the most significantly correlated variable to the model. This step is repeated until all significant covariates are included in the model.

We implemented both backward and forward selection models setting a threshold p-value of 0.10 for what is considered to be significant, and thus selected in the model. Thus, the final set of selected variables was based on whether they were chosen in either a stepwise selection approach or if they were selected based on our theoretical understanding of key covariates that should be correlated with the outcome of interest. The full list of covariates used in our preferred fixed effects estimation is presented in Table 25 below. Note that child-level covariates are only used in regressions estimating impact at the child level (i.e. where the dependent variable is defined at the child level). The rest of the covariates are used in all regressions used to estimate the main results presented in Volume 1.

In our preferred fixed effects estimation, we allow these covariates to vary over time given that a household-level fixed effects estimation already accounts for (observed or unobserved) time-invariant characteristics at the household or higher level. By including observed time-varying characteristics in our fixed effects regressions we make it possible to control for additional relevant confounding factors.

Table 25 Covariates used in our preferred fixed effects regression model

Level	Definition
Community	
	Community suffered from floods in the last 12 months
Household	
	Household owns cows
	Household's roof material is cardboard
	Household hunger scale score is 1: little to no household hunger
	Household does not have a roof
	Household head has any Islamia education
	Household hunger scale score is 2: moderate household hunger
	Household head has started but not completed primary school
	Household travel time to the closest place to buy malaria medicine is 0 to less than 30 mins
	Household's main cooking fuel is straw/shrubs/grass
	Household's roof material is mud/mud bricks
	Main economic activity of the household head is farming/herding for commercial sale
	Household's wall material is earth
	Household owns a satellite television
	Household owns a mobile phone
	Household's floor material is cement/concrete
	Household's wall material is bamboo with mud

	Household's poverty score
	Household owns a boat
	Household's roof material is sod
	Household's roof material is roofing shingles
	Household owns an electric iron
	Household's wall material is other
	Main economic activity of the household head is owning his/her own business
	Household's rood material is thatch
	Main economic activity of the household head is paid employment
	Number of children under 18 in the household
	Household's wall material is stone/bricks/cement blocks/cement
	Household size
	Household has electricity
	Household's wall material is stone with mud
	Household's wall material is cane/palm/trunks
	Household's wall material is adobe (sun baked bricks)
	Number of household members 18 or above in the household
	Household travel time to the closest health facility is two hours to less than five hours
	Household is using improved toilet (three types)
	Household's main water source is located outside the compound or further away
	Household is using improved toilet
Mother	
	Mother has no education
	Mother went to primary school
	Mother went to secondary school
	Mother is married (monogamous)
	Mother is married (polygamous)
	Mother is divorced/separated
	Mother was never married
	Mother is widowed
	Age of the mother
	Mother has a spouse
Child	
	Age of the child in months
	Sex of the child

3.3.2 Main results and robustness checks

As described above, we ran different regression specifications to check the consistency and robustness of our results with respect to our main regression specification explained above (and reported in Volume 1).

Table 26 to Table 30 present the results of 11 different specifications for a subset of key breastfeeding and undernutrition outcomes. (Further regression results from other outcome indicators can be obtained upon request.) In the majority of cases, our conclusions are robust to changes in the regression specification. In Volume 1 we explicitly mention where evidence is slightly less robust to such changes.

In general, the tables begin by presenting results from a simple, reduced-form regression (1), and then present results from models that continuously build up in terms of complexity, in order to understand the effects of different sets of controls and key variations on our models.

Specifications (1) to (6) are estimated using OLS and a DID model with state and LGA dummies. These models take into account the sampling structure of the data by including sampling weights, stratification and clustering of standard errors. The inclusion of sampling weights and stratification at the LGA level provides more efficient estimators (i.e. with reduced variance). The implications of clustering are discussed in Section 3.3.1 above.

Specifications (7) to (8) are estimated using fixed effects and a DID model with household fixed effects. These models also take into account the sampling structure of the data by including sampling weights and clustering of standard errors. Stratification allows subgroups (i.e. LGAs in our sampling design) to have different time-invariant distribution of the error terms. Thus, fixed effects models do not allow us to control for stratification in the sample design, given that these models can only estimate coefficients for variables that change over time. However, the advantage of fixed effects models is that they go further in controlling also for time-invariant unobserved characteristics at the household level, in addition to LGA and state level.

To be more precise:

- Specification (1) starts with an OLS DID model with no additional controls, i.e. only including the Post and Treat terms, and the interaction between these two, as specified in the equation above. This specification also includes sampling weights corrected for attrition (see Section 3.1.3 on attrition for an explanation of the methodology).
- Specification (2) estimates the same model but includes sampling weights that are not corrected for attrition. Note that for each outcome, the estimated treatment effect (DID) and its significance is stable to changes in the sampling weights definition. We include this specification to show that adding attrition correction to weights in this impact evaluation does not have a significant effect on our estimates.
- Specifications (3) and (4) add controls to (1). Control variables are selected using the same process described above: that is, a combination of theoretical considerations and employing the stepwise optimisation algorithm (see discussion of choice of covariates in Section 3.3.1). Final covariates include a similar subset of individual, household, and village-level characteristics, as described in Table 25. Given that we are operating in a household panel setting, we prefer to implement OLS DID models using baseline levels of covariates to avoid including variables that might have been affected by the programme over time, which would introduce endogeneity into our model. Thus, all household and village-level covariates used in these models are set at the baseline level. However, we still use individual-level (child and mother) characteristics that vary over time since we interview different individuals over time within the same households, i.e. because we do not have a panel of children or mothers at hand.
- Specification (4) includes the same controls as in (3) but adds in the distance to the nearest WINNN-supported health facility, given that geographical location could be a potential confounder correlated to both the outcome and the access to treatment. In the case of undernutrition indicators, it also includes dummy variables for the anthropometric specialist in order to control for the potential differential measurement error between baseline and endline identified in the data quality assessment (see Section 3.2), since this could be correlated with such individual specialists. These specifications are included to show how controlling for such specific confounding factors in our regression changes point estimates for treatment effects on different outcome indicators. Among OLS estimations, these are our preferred specifications.

- Specifications (5) and (6) are similar to (3) and (4), respectively, but they now cluster standard errors at the LGA level as opposed to the EA (see discussion of clustering of standard errors in Section above for further details regarding the implications of different levels of clustering). Note that point estimates of these regressions generally remain the same as in (3) and (4), but significance levels change, given that standard errors are now different. We include these specifications here to show the effects of changing the level of clustering on our impact estimates.
- In specifications (7) to (11) we use a fixed effects estimation procedure to estimate treatment effects, controlling for household fixed effects (see details on the fixed effects specification in Section 0 above). Specification (7) starts with a fixed effects model with no additional controls and includes sampling weights that are not corrected for attrition. Specification (8) uses a different definition of weights compared to (7). Estimating fixed effects 3.3.1 models does not allow us to include time-varying weights. For specification (8) we therefore use attrition-corrected weights at both baseline and endline, restricting the model to households that appear in both stages of analysis (hence the loss of observations). This does not correct for the attrition problem fully but is a reasonable approximation given the restrictions of the fixed effects model. For all outcomes, coefficient and significance levels of the treatment effects are fairly robust to changes in the weights definition, showing only slight variations in some cases. Specifications (7) and (8) cluster standard errors at the household level.
- Specifications (9) to (11) are similar to (7), but they add additional controls. Controls in the fixed effects model setting include a subset of individual, household, and village-level time-varying characteristics selected through a combination of theoretical considerations and stepwise first-stage regressions. We cannot include baseline levels of these characteristics as in the OLS DID models since characteristics (observable and unobservable) that are constant across time are controlled for by including household fixed effects here (see discussion of the choice of covariates in Section 3.3.1).
- Additionally, as in the OLS DID models, we also change the way standard errors are calculated in the fixed effects regressions. Column (9) clusters at the level of the household, column (10) clusters at the level of the PSU that is at the EA level, and column (11) clusters at the treatment level that is at the LGA level. In general, estimates are robust to changes in the level of clustering, except for a few cases[3] where the level of significance is lost when we include clustering at the LGA level.

In general, we find that estimates derived from fixed effects regressions (7) to (11) are not significantly different to estimates derived from our preferred OLS DID estimations, such as, for example, specification (4). Some outcomes present slight variations in the point estimates of the treatment effect and their level of significance but this is expected since the fixed effects models also account for unobserved characteristics at the household level that do not change over time, which DID models in specifications (1) to (6) do not.

Specification (10) was chosen as our preferable model since it includes household-level fixed effects and controls for observed characteristics that might change over time, thereby controlling for both observable and unobservable confounding factors that might be related to treatment assignment and outcome indicators. This model also clusters standard errors at the level of the PSU or EA level, which is our preferred model, given the reasons outlined in the discussion of the clustering of standard errors in Section 3.3.1 above. This model does not explicitly correct for attrition, but results in specification (8) show that attempting to do so does not

[3] Impact estimates of the following indicators lose statistical significance when changing the level of clustering from EA to LGA in the fixed effects models: early initiation in breastfeeding, knowledge of mothers about standard feeding times, and DPT/penta vaccination.

change our conclusions and leads to a loss in observations included in the estimation. The outputs from specification (10) are therefore presented in Volume 1 of this report.

Table 26 Regressions: Early initiation (<24 hours) to breastfeeding (children 0–23 months)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Treatment*Post (DID)	**0.059***	**0.062***	**0.063***	**0.058***	**0.063**	**0.058**	**0.075****	**0.070***	**0.093****	**0.093****	**0.093**
	(0.030)	**(0.029)**	**(0.028)**	**(0.028)**	**(0.050)**	**(0.052)**	**(0.028)**	**(0.029)**	**(0.030)**	**(0.032)**	**(0.052)**
Treatment	0.121	0.128*	0.087	0.080	0.087**	0.080**					
	(0.064)	(0.064)	(0.057)	(0.060)	(0.024)	(0.027)					
Time (BL=0, EL=1)	0.126***	0.128***	0.123***	0.123***	0.123***	0.123***	0.129***	0.130***	0.109***	0.109***	0.109**
	(0.024)	(0.023)	(0.022)	(0.022)	(0.032)	(0.032)	(0.021)	(0.022)	(0.024)	(0.027)	(0.033)
Constant	0.603***	0.596***	0.258	0.244	0.258	0.244	0.616***	0.621***	0.439*	-0.107	-0.107
	(0.069)	(0.069)	(0.168)	(0.168)	(0.142)	(0.135)	(0.007)	(0.007)	(0.188)	(0.432)	(0.391)
Weights	Attrition-corrected	Non-attrition-corrected	Attrition-corrected	Attrition-corrected	Attrition-corrected	Attrition-corrected	Non-attrition-corrected	Attrition-adjusted weights for BL and EL	Non-attrition-corrected	Non-attrition-corrected	Non-attrition-corrected
Level of clustering of standard error	EA	EA	EA	EA	LGA	LGA	Household	Household	Household	EA	LGA
Stratification	LGA	LGA	LGA	LGA	LGA	LGA	n.a	n.a	n.a	n.a	n.a
State dummies	Yes	Yes	Yes	Yes	Yes	Yes	n.a	n.a	n.a	n.a	n.a
LGA dummies	Yes	Yes	Yes	Yes	Yes	Yes	n.a	n.a	n.a	n.a	n.a
Household fixed effects	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes
Controls	No	No	Baseline level	Distance + baseline level	Baseline level	Distance + baseline level	No	No	Time-varying covariates	Time-varying covariates	Time-varying covariates
Observations	7,940	8,118	7,459	7,406	7,459	7,406	8,118	7,607	7,592	7,592	7,592
R-squared	0.090	0.089	0.162	0.162	0.162	0.162	0.056	0.057	0.088	0.088	0.088
Number of households							3,045	2,818	2,983	2,983	2,983
Standard errors in parentheses											
***** p<0.001,** p<0.01,* p<0.05**											

Table 27 Regressions: Exclusive breastfeeding (children 0–5 months)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
1.treatment#1.post	**0.058***	**0.056***	**0.063***	**0.062***	**0.063**	**0.062**	**0.021**	**0.039**	**0.031**	**0.031**	**0.031**
	(0.028)	**(0.028)**	**(0.029)**	**(0.029)**	**(0.043)**	**(0.043)**	**(0.050)**	**(0.054)**	**(0.052)**	**(0.053)**	**(0.042)**
1.treatment	-0.010	-0.009	0.021	0.037	0.021	0.037					
	(0.027)	(0.027)	(0.040)	(0.043)	(0.029)	(0.029)					
1.post	0.042*	0.044**	0.037*	0.037*	0.037	0.037	0.040	0.021	0.081*	0.081*	0.081
	(0.017)	(0.017)	(0.017)	(0.017)	(0.025)	(0.025)	(0.030)	(0.032)	(0.037)	(0.036)	(0.043)
Constant	0.115*	0.109*	0.049	0.036	0.049	0.036	0.069***	0.077***	-0.349	-0.294	-0.294
	(0.055)	(0.053)	(0.162)	(0.162)	(0.158)	(0.158)	(0.011)	(0.012)	(0.245)	(0.244)	(0.279)
Weights	Attrition-corrected	Non-attrition-corrected	Attrition-corrected	Attrition-corrected	Attrition-corrected	Attrition-corrected	Non-attrition-corrected	Attrition-adjusted weights for baseline and endline	Non-attrition-corrected	Non-attrition-corrected	Non-attrition-corrected
Level of clustering of standard error	EA	EA	EA	EA	LGA	LGA	Household	Household	Household	EA	LGA
Stratification	LGA	LGA	LGA	LGA	LGA	LGA	n.a	n.a	n.a	n.a	n.a
State dummies	Yes	Yes	Yes	Yes	Yes	Yes	n.a	n.a	n.a	n.a	n.a
LGA dummies	Yes	Yes	Yes	Yes	Yes	Yes	n.a	n.a	n.a	n.a	n.a
Household fixed effects	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes
Controls	No	No	Baseline level	Distance + baseline level	Baseline level	Distance + baseline level	No	No	Time-varying covariates	Time-varying covariates	Time-varying covariates
Observations	2,031	2,080	1,911	1,900	1,911	1,900	2,080	1,954	1,948	1,948	1,948
R-squared	0.071	0.072	0.129	0.131	0.129	0.131	0.011	0.008	0.195	0.195	0.195
Number of households							1,470	1,381	1,407	1,407	1,407
Standard errors in parentheses											
***** p<0.001, ** p<0.01, * p<0.05**											

Table 28 Regressions: Proportion of children who are wasted (6–35 months)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Treatment*Post (DID)	**0.031**	**0.030**	**0.019**	**0.019**	**0.019**	**0.019**	**0.036**	**0.042***	**0.004**	**0.004**	**0.004**
	(0.019)	**(0.019)**	**(0.019)**	**(0.021)**	**(0.019)**	**(0.009)**	**(0.020)**	**(0.020)**	**(0.020)**	**(0.021)**	**(0.018)**
Treatment	-0.069*	-0.066*	-0.064*	-0.043	-0.064**	-0.043*					
	(0.028)	(0.029)	(0.031)	(0.034)	(0.018)	(0.017)					
Time (BL=0, EL=1)	-0.003	-0.001	-0.001	0.034	-0.001	0.034	-0.010	-0.014	0.005	0.005	0.005
	(0.013)	(0.013)	(0.014)	(0.051)	(0.011)	(0.048)	(0.015)	(0.015)	(0.018)	(0.018)	(0.014)
Constant	0.203***	0.199***	0.373**	0.343*	0.373**	0.343*	0.165***	0.165***	-0.088	-0.261	-0.261
	(0.036)	(0.036)	(0.130)	(0.138)	(0.124)	(0.145)	(0.005)	(0.005)	(0.190)	(0.207)	(0.209)
Weights	Attrition-corrected	Non-attrition-corrected	Attrition-corrected	Attrition-corrected	Attrition-corrected	Attrition-corrected	Non-attrition-corrected	Attrition-adjusted weights for BL and EL	Non-attrition-corrected	Non-attrition-corrected	Non-attrition-corrected
Level of clustering of standard error	EA	EA	EA	EA	LGA	LGA	Household	Household	Household	EA	LGA
Stratification	LGA	LGA	LGA	LGA	LGA	LGA	n.a	n.a	n.a	n.a	n.a
State dummies	Yes	Yes	Yes	Yes	Yes	Yes	n.a	n.a	n.a	n.a	n.a
LGA dummies	Yes	Yes	Yes	Yes	Yes	Yes	n.a	n.a	n.a	n.a	n.a
Household fixed effects	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes
Controls	No	No	Baseline level	Anthropometric specialist + distance + baseline level	Baseline level	Anthropometric specialist + distance + baseline level	No	No	Time-varying covariates	Time-varying covariates	Time-varying covariates
Observations	9,600	9,810	8,968	8,858	8,968	8,858	9,810	9,167	9,122	9,122	9,122
R-squared	0.008	0.008	0.050	0.059	0.050	0.059	0.001	0.001	0.045	0.045	0.045
Number of households							3,266	3,004	3,200	3,200	3,200
Standard errors in parentheses											
***** p<0.001, ** p<0.01, * p<0.05**											

Table 29 Regressions: Proportion of children who are stunted (0–35 months)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Treatment*Post (DID)	**-0.010**	**-0.006**	**-0.007**	**-0.018**	**-0.007**	**-0.018**	**-0.003**	**-0.008**	**0.002**	**0.002**	**0.002**
	(0.025)	**(0.025)**	**(0.024)**	**(0.026)**	**(0.047)**	**(0.016)**	**(0.024)**	**(0.025)**	**(0.024)**	**(0.027)**	**(0.044)**
Treatment	0.013	0.013	-0.034	-0.049	-0.034	-0.049*					
	(0.046)	(0.042)	(0.039)	(0.039)	(0.022)	(0.018)					
Time (BL=0, EL=1)	-0.020	-0.022	-0.020	-0.153**	-0.020	-0.153**	-0.022	-0.016	-0.015	-0.015	-0.015
	(0.016)	(0.016)	(0.015)	(0.047)	(0.035)	(0.043)	(0.017)	(0.018)	(0.020)	(0.020)	(0.038)
Constant	0.566***	0.570***	0.179	0.314*	0.179	0.314**	0.533***	0.528***	0.314	0.055	0.055
	(0.051)	(0.047)	(0.138)	(0.137)	(0.110)	(0.100)	(0.005)	(0.006)	(0.189)	(0.178)	(0.178)
Weights	Attrition-corrected	Non-attrition-corrected	Attrition-corrected	Attrition-corrected	Attrition-corrected	Attrition-corrected	Non-attrition-corrected	Attrition-adjusted weights for BL and EL	Non-attrition-corrected	Non-attrition-corrected	Non-attrition-corrected
Level of clustering of standard error	EA	EA	EA	EA	LGA	LGA	Household	Household	Household	EA	LGA
Stratification	LGA	LGA	LGA	LGA	LGA	LGA	n.a	n.a	n.a	n.a	n.a
State dummies	Yes	Yes	Yes	Yes	Yes	Yes	n.a	n.a	n.a	n.a	n.a
LGA dummies	Yes	Yes	Yes	Yes	Yes	Yes	n.a	n.a	n.a	n.a	n.a
Household fixed effects	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes
Controls	No	No	Baseline level	Anthropometric specialist + distance + baseline level	Baseline level	Anthropometric specialist + distance + baseline level	No	No	Time-varying covariates	Time-varying covariates	Time-varying covariates
Observations	11,570	11,829	10,833	10,707	10,833	10,707	11,829	11,065	11,025	11,025	11,025
R-squared	0.023	0.022	0.170	0.192	0.170	0.192	0.001	0.001	0.164	0.164	0.164
Number of households							3,345	3,072	3,289	3,289	3,289
Standard errors in parentheses											
***** p<0.001, ** p<0.01, * p<0.05**											

Table 30 Regressions: Proportion of children who are underweight (0–35 months)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Treatment*Post (DID)	**0.008**	**0.010**	**0.007**	**0.005**	**0.007**	**0.005**	**0.012**	**0.015**	**0.009**	**0.009**	**0.009**
	(0.023)	**(0.023)**	**(0.023)**	**(0.028)**	**(0.046)**	**(0.013)**	**(0.023)**	**(0.024)**	**(0.026)**	**(0.027)**	**(0.044)**
Treatment	-0.101**	-0.087*	-0.115**	-0.123**	-0.115***	-0.123***					
	(0.038)	(0.036)	(0.039)	(0.044)	(0.021)	(0.023)					
Time (BL=0, EL=1)	-0.012	-0.010	-0.015	-0.158**	-0.015	-0.158***	-0.005	-0.011	-0.016	-0.016	-0.016
	(0.017)	(0.017)	(0.017)	(0.050)	(0.040)	(0.030)	(0.016)	(0.017)	(0.020)	(0.022)	(0.043)
Constant	0.523***	0.508***	0.475**	0.564***	0.475**	0.564***	0.387***	0.386***	0.190	-0.037	-0.037
	(0.043)	(0.041)	(0.151)	(0.158)	(0.141)	(0.139)	(0.005)	(0.006)	(0.187)	(0.181)	(0.144)
Weights	Attrition-corrected	Non-attrition-corrected	Attrition-corrected	Attrition-corrected	Attrition-corrected	Attrition-corrected	Non-attrition-corrected	Attrition-adjusted weights for baseline and endline	Non-attrition-corrected	Non-attrition-corrected	Non-attrition-corrected
Level of clustering of standard error	EA	EA	EA	EA	LGA	LGA	Household	Household	Household	EA	LGA
Stratification	LGA	LGA	LGA	LGA	LGA	LGA	n.a	n.a	n.a	n.a	n.a
State dummies	Yes	Yes	Yes	Yes	Yes	Yes	n.a	n.a	n.a	n.a	n.a
LGA dummies	Yes	Yes	Yes	Yes	Yes	Yes	n.a	n.a	n.a	n.a	n.a
Household fixed effects	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes
Controls	No	No	Baseline level	Anthropometric specialist + distance + baseline level	Baseline level	Anthropometric specialist + distance + baseline level	No	No	Time-varying covariates	Time-varying covariates	Time-varying covariates
Observations	11,742	12,002	10,983	10,857	10,983	10,857	12,002	11,225	11,173	11,173	11,173
R-squared	0.033	0.031	0.070	0.093	0.070	0.093	0.000	0.000	0.035	0.035	0.035
Number of households							3,352	3,076	3,300	3,300	3,300
Standard errors in parentheses											
***** p<0.001, ** p<0.01, * p<0.05**											

As described above, all our analyses aim to accurately take into account the fact that we are dealing with sampled survey data and that estimates should be representative for the areas surveyed. The methodology for calculating sampling weights at baseline is described in Annex C.4. In addition, all point estimates of descriptive indicators presented in Volume 1 include sampling weights corrected for attrition (see Section 3.1.3 on attrition and Section 3.1.1 on sampling weights). Our impact estimates presented in Volume 1, however, are derived from fixed effects estimations that include sampling weights but no attrition correction, given that fixed effects estimations do not allow for time-varying weights to be included in the analysis. We further test whether this is an issue in specification (10) and, comparing this to (9), conclude that the risk of bias is low given that inferences about the effects of WINNN support do not change. (See above for the specific methodological differences between specifications (10) and (9).)

3.3.3 Threats to the DID and other limitations

As mentioned in Volume 1, a DID approach provides unbiased treatment effects if the difference in outcomes between treatment and control areas would have remained the same over time in the absence of the WINNN intervention. Thus, it is important to analyse the risk posed by differential contamination or spillover effects in treatment and control areas over the study period, which would violate that assumption and lead to biased impact estimates. Section 2.3.3 in Volume 1 (Quantitative Impact Evaluation of the WINNN Programme – Volume 1, 2017) addresses these issues in detail, and concludes that the potential for bias due to contamination or spillover effects is low and not problematic for our overall impact estimates.

In this section, we provide evidence from a geographical analysis that shows the existence of spillovers in relation to accessing community IYCF counselling, i.e. individuals in control LGAs accessing WINNN-supported IYCF interventions in treatment areas. This suggests that, although not problematic for our impact estimates (as explained in Volume 1), spillovers across LGAs did materialise. Also, secondary evidence collected throughout the implementation of this evaluation indicates that no programmes that focused on IYCF promotion were implemented in control areas. Thus, we reject the possibility of there being a contamination issue.

As explained in Volume 1, the risk of bias of treatment effects resulting from spillover across LGAs is low and not problematic. In this section we also explore whether spillovers of the IYCF interventions and the CMAM programme could have materialised differently by states. We conclude that the risk of bias due to spillover effects remains low at the state level.

Geographical analysis of spillover effects

Table 31 presents the change from baseline to endline in control areas in regard to the access to IYCF and CMAM interventions by different measures of distance to treatment units. The first measure used – close versus far distance – measures the geographical distance from a control LGA to the nearest treatment facility (facilities providing IYCF-related services or CMAM facilities) in a WINNN LGA. Individuals in control communities with distance values lower than the median are part of the close distance group (column 1 in Table 31 below). Analogously, the far distance group includes individuals with distance values larger than the median (column 2).

The second measure – neighbour versus non-neighbour– identifies whether a control LGA is geographically adjacent to a treatment LGA. Individuals living in control communities which are adjacent to treatment ones are categorised in the neighbour group (column 3), while individuals living in non-adjacent communities are in the non-neighbour group (column 4).

Table 31 shows that the change over time in control areas (time variable) is statistically significant at a 95% confidence level or more only for individuals in the close distance and neighbour groups and for the indicator on community IYCF exposure to community IYCF counselling. This suggests that mothers in control areas living close to treatment units increased their exposure to community IYCF counselling due to the WINNN intervention. This provides evidence of spillovers in the exposure to community IYCF counselling.

Table 31 Spillover effects of community IYCF counselling and MUAC measurement in control areas

	Proportion of mothers (15–49 years) who ever attended IYCF counselling in the community				Proportion of children who have ever had their MUAC measured (children 6–35 months)			
	Close distance	Far distance	Neighbour	Non-neighbour	Close distance	Far distance	Neighbour	Non-neighbour
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Time (BL=0, EL=1)	**0.054*** **	**0.023**	**0.041** **	**0.044***	**0.022**	**-0.015**	**0.015**	**-0.004**
	(0.020)	**(0.021)**	**(0.019)**	**(0.023)**	**(0.018)**	**(0.016)**	**(0.017)**	**(0.019)**
Constant	0.458**	0.010	0.236	-0.121	0.092	0.346***	0.195***	0.354**
	(0.218)	(0.099)	(0.159)	(0.162)	(0.104)	(0.105)	(0.071)	(0.153)
Observations	2,488	2,486	3,037	1,937	2,440	2,455	2,981	1,914
Number of households	833	819	965	687	824	806	953	677
Adjusted R-squared	0.077	0.027	0.067	0.044	0.019	0.023	0.011	0.024

Notes: Regressions only include control LGAs. Significance levels are reported with stars *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are reported in parentheses. Regressions include household fixed effects and household and individual-level covariates.

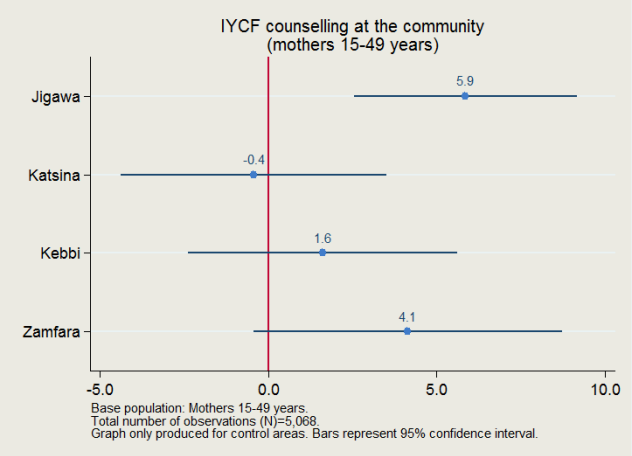
State analysis of spillovers effects

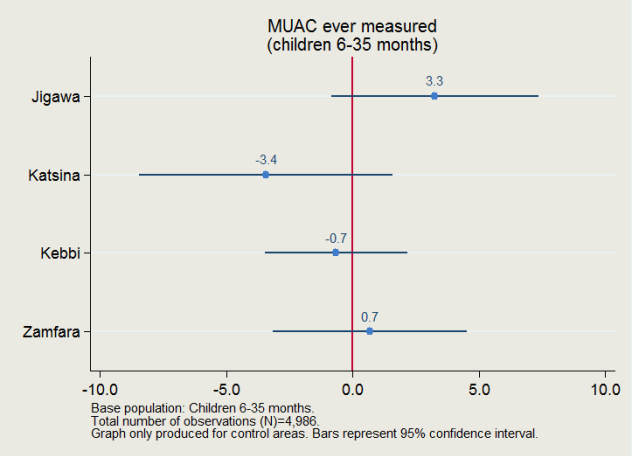
We also investigate whether spillovers of IYCF and CMAM interventions could have materialised differently by states. For this we analyse how exposure to community-IYCF counselling and undernutrition screening changed over time in control areas across the four states that form part of this evaluation. Figure 12 plots the estimated differences across time in those indicators in control areas and include 95% confidence intervals.

Results show that the exposure to community IYCF counselling in control areas increased significantly between baseline and endline only in Jigawa. This indicates that the impact on community IYCF counselling exposure in Jigawa could have been larger had spillovers not taken place, which could mean that we potentially underestimate treatment effects of WINNN supported interventions for that state. However, further analysis presented in Table 12 in Section 4.1.2 of Volume 1 report (Quantitative Impact Evaluation of the WINNN Programme – Volume 1, 2017) show that the risk of such potential downward bias is likely to be low given that a significant statistical impact on the exposure to community IYCF counselling is consistently found in treatment areas across all states.

In the case of the CMAM programme, results presented in Figure 12 indicate that spillovers in terms of MUAC measurement among children could be an issue in Jigawa since that state presents the largest change in the proportion of children in control areas being screened for undernutrition between baseline and endline, however that difference is not statistically significant. This evidence suggests that the risk of an underestimation of impacts influenced by the CMAM programme due to spillover effects is also likely to be low.

Figure 12 Change over time in control areas for IYCF counselling attendance and MUAC measurement by states





3.4 Supplementary PSM analysis

3.4.1 Aim

As described in Section 3.3.1 above, the main impact estimates presented in this quantitative evaluation are estimates of ITT effects, i.e. estimates of the average effect of WINNN across the population of individuals in WINNN-supported LGAs, irrespective of whether these individuals have actually been in touch with or have heard of any of the WINNN components or not. This means the estimated effects of WINNN-supported activities can be 'diluted' by the fact that individuals have not actually been in touch with WINNN. Getting around this analytical issue is complicated for WINNN overall, given that it is a complex, multi-pronged programme. However, it is possible to try to estimate different treatment effects for sub-components supported by WINNN.

In Box 1 in Volume 1 of this report (Quantitative Impact Evaluation of the WINNN Programme – Volume 1, 2017), **and following an explicit request from stakeholders involved in this evaluation,** we therefore present results from a supplementary PSM analysis, where we try to estimate the direct average treatment effects on the treated (ATT) of IYCF counselling of caregivers on exclusive breastfeeding rates among children aged 0–5 months, **using endline data only**. Note that this was a supplementary exploratory analysis only, and more complex modelling that would have taken into account baseline data and matching across time was not feasible. Caveats resulting from this limitation are presented below.

It is important to reiterate here that we in fact implement three separate PSM analyses here. These are three treatment effect estimations, of which one can be interpreted as a robustness analysis:

- First, comparing children with mothers who did attend IYCF counselling in **WINNN-supported LGAs** with children whose mothers did not attend IYCF counselling in **non-supported LGAs. (Treatment effect 1, ATT1)**. This is, in effect, comparing treated observations from treatment LGAs with comparison observations from non-treatment LGAs.
- Second, comparing children with mothers **who did attend IYCF counselling in WINNN-supported LGAs** with children whose mothers did not attend IYCF counselling **in WINNN-supported LGAs. (Treatment effect 2, ATT2)**. This is a comparison of treated observations with comparison observations within treatment LGAs. The difference between ATT1 and ATT2 is that, if we assume that WINNN has an effect on exclusive breastfeeding decisions beyond

direct IYCF counselling in WINNN-supported LGAs, ATT1 will be the sum of both the direct treatment effect of IYCF counselling and such additional 'WINNN effects'. ATT2, on the other hand, will be an estimate of the direct effects of the IYCF interventions only, because the comparison group, taken from within the WINNN-supported LGA, will also have benefited from additional 'WINNN effects'. We therefore assume ATT1 to be larger than ATT2.

- Finally, as robustness analysis, and to see whether 'WINNN effects' actually materialise, we implement a third PSM analysis, **comparing children with mothers who never attended IYCF counselling in WINNN-supported LGAs to children with mothers who never attended IYCF counselling in control LGAs (Treatment effect 3, ATT3).** Any significant differences between these groups should be due to the 'living in WINNN LGAs' effect, and not due to direct IYCF counselling. We expect this estimate to roughly correspond to the difference between ATT1 and ATT2.

Using the potential outcomes framework, these three effects can be expressed as follows:

$$(\mathbf{2})\ \mathbf{ATT2} = \mathrm{E}(\mathrm{Y}_{[\mathrm{IYCF}=1,\mathrm{T}=1]} - \mathrm{Y}_{[\mathrm{IYCF}=0,\mathrm{T}=1]}),$$
$$(\mathbf{3})\ \mathbf{ATT3} = \mathrm{E}(\mathrm{Y}_{[\mathrm{IYCF}=0,\mathrm{T}=1]} - \mathrm{Y}_{[\mathrm{IYCF}=0,\mathrm{T}=0]}),$$
$$(\mathbf{4})\ \mathbf{ATT1} = \mathrm{E}(\mathrm{Y}_{[\mathrm{IYCF}=1,\mathrm{T}=1]} - \mathrm{Y}_{[\mathrm{IYCF}=0,\mathrm{T}=0])} = \mathrm{ATT2} + \mathrm{ATT3},$$

Where the subscripts indicate treatment with IYCF counselling ($\mathrm{IYCF} = 1$) or not ($\mathrm{IYCF} = 0$), and whether the individual lives in a WINNN-supported LGA ($\mathrm{T} = 1$) or not ($\mathrm{T} = 0$). The outcome of interest is expressed by Y: in the present case, exclusive breastfeeding. ATT2 corresponds to the expected difference in outcomes within WINNN-supported LGAs between observations with and without IYCF counselling, whereas ATT3 corresponds to the difference in outcomes between WINNN-supported and non-supported LGAs among observations that do not receive IYCF counselling. This set of equations makes clear that in fact the overall treatment effect that we are mainly interested in, ATT1, is composed of ATT2 and ATT3, i.e. the direct effect of IYCF counselling plus the effect of being in the WINNN-supported LGA. We estimate all three treatment effects using PSM. Note that the equations above also show that, depending on which ATT estimate we look at, treatment and control groups in the present context are defined slightly differently.

3.4.2 Estimation strategy

The key problem that PSM attempts to solve is selection bias. In the present case, this problem arises because both IYCF counselling and WINNN support in LGAs has not been randomly allocated to mothers, which means that there are potentially systematic differences between children whose mothers did receive IYCF counselling, children whose mothers did not receive IYCF counselling, children who are in WINNN- supported LGAs, and children in non-supported LGAs. Such systematic differences could plausibly be related to outcome measures that this evaluation is interested in. For example, it could be that mothers with a higher education status are both more likely to go to IYCF counselling and to exclusively breastfeed their children. This means that there would be a systematic difference in terms of education status among mothers who did and did not receive IYCF counselling that is also related to the outcome of interest.

This in turn implies that observed dissimilarities in outcome measures across individuals from treatment and control groups could be due to underlying systematic differences and not IYCF counselling or WINNN activities in WINNN-supported LGAs, i.e. 'WINNN effects', itself/themselves. Simple comparisons of outcome indicators across such groups would be invalid and biased, as regards inferring programme impact, because these groups cannot be assumed to be alike. This is the problem of selection bias.

PSM tackles this problem by using data from the control group to construct appropriate comparisons with individuals in the treatment group, thus building a valid counterfactual. This happens by matching and comparing outcomes for units in the treatment group with control units that are as similar as possible to each other according to a set of relevant observable characteristics, i.e. comparing like with like only. Relevant characteristics are those that are thought to be driving selection bias. These are the characteristics that are systematically different across treatment status groups and that are related to outcome measures of interest. When appropriately controlling for all of these characteristics, selection bias is also controlled for.

Specifically, PSM is a two-stage analytical approach that employs a propensity score as a 'comparator metric' that summarises the information about the set of relevant characteristics, i.e. the ones that drive selection bias, defined above. The first stage of any PSM analysis is to compute a valid propensity score for each unit of observation. The second stage is to then compare outcome indicators of interest across units (i.e. children, in this case) with similar propensity scores.

PSM first-stage model selection

To estimate the propensity score in the first stage, and for each of the estimations of ATT1, ATT2, and ATT3, this study followed the procedure suggested by Imbens and Rubin (2015, p. 281 ff.). The underlying model specification for this procedure is either a logit or probit regression for the first stage. This means that the propensity scores are estimated by first specifying treatment and control assignment as a binary variable that has the values 0 (for control) and 1 (for treatment). The estimated scores are then modelled as the fitted values that are derived from a logit or probit estimation, with the binary treatment variables as dependent variable and the covariates across which balance is supposed to be achieved as the regressors. These fitted values lie between 0 and 1.

To be more concrete, in the case of a logistic regression specification, the binary response variable is modelled as follows:

$$(5)\quad \Pr(T = 1 \mid X_i) = \frac{e^{f(X_i)}}{1+e^{f(X_i)}},$$

Where $\Pr(T = 1 \mid X_i)$ is the probability of the treatment indicator (T) being equal to one, conditional on the covariates (X_i) for unit i. The function $f(X)$ is normally modelled linearly, i.e. is of the form $f(X) = X\beta$. The coefficients of this function (β) are estimated using maximum likelihood techniques. The fitted values, i.e. the predicted probabilities that follow from this procedure, are the propensity scores for each unit of observation.

The key question for the first stage is which covariates to include in $f(X)$ so that this procedure produces a valid estimate of the propensity score. Building on the procedure described in Imbens and Rubin (2015) for selecting covariates, this study implemented a three-step approach to make this decision:

1. **Select a set of basic covariates based on substantive grounds**:

   The starting point for the PSM analysis was to select variables that were likely to be relevant and valid for this analysis from a theoretical perspective. 'Relevant' in this case meant that variables had to be selected that were theoretically expected to be very likely to be correlated with treatment status and treatment effects, thereby introducing selection bias in a simple comparison of treatment outcomes between control and treatment groups.

'Valid' in this case meant that variables had to be selected that were expected to not be influenced by the programme. This is because this PSM analysis was implemented using endline data only. At baseline, this would not be an issue, given that the programme had not started yet and no variable would have been influenced by the programme. At endline, however, this was not the case. Hence, only variables were used for PSM for which a plausible argument could be made that they have not been influenced by the programme at endline. As explained above, note that due to the limited scope of this analysis, more complex modelling that would have taken into account baseline data as well was not feasible.

2. **Increase the set of valid covariates based on algorithmic approaches:**

   In addition, this study employed variable selection algorithms to identify valid variables, i.e. variables that were not affected by the programme, and that were significantly correlated both with the treatment status and the outcome variable. There are a variety of methods available to do this. This study's approach was to implement stepwise regressions. Such regressions are commonly used and easily implemented algorithms to select independent variables based on significant correlations with certain dependent variables.

   There are two stepwise regression approaches that can be employed for this: backward and forward stepwise regression. The underlying idea behind both approaches is to check each covariate, step-by-step, for significant correlation with the outcome and treatment assignment variable separately. Such a correlation is relevant because variables that possibly bias impact estimates will have some relation to both the treatment status and the outcome looked at.

   Backward selection starts with the full set of covariates, i.e. a regression including all variables, and then discards the term that is least significantly correlated with the dependent variable. It continues to do so until all variables that are uncorrelated with the dependent variable are discarded. Forward selection, instead, starts with an empty set of covariates, i.e. a regression on a constant, and then checks the significance of each covariate if it is included in the regression. It then adds the most significantly correlated variable to the model. This step is repeated until all significant covariates are included in the model.

   Both for backward and forward estimation a threshold p-value for what is considered to be significant needs to be specified. For backward selection, this means setting the level for identifying whether all variables that are uncorrelated with the outcome variable have been discarded: if the p-value of the least significant variable remaining is under the threshold, i.e. all the variables still included in the model are even more significant, the procedure stops. For forward selection, this means setting the level for identifying whether all significant covariates have been included in the model: if the p-value of the most significant variable to be added is equal to the threshold, i.e. the significance level of all variables that have not yet been included in the model is equal to or below the threshold, the procedure stops. Setting this threshold therefore influences the variables that are selected in stepwise regressions.

   This study implemented both backward and forward selection, using baseline data and using thresholds of $p = 0.05$. The analysis employs this covariate selection procedure on both the outcome variable and the relevant treatment status, given the importance of determining the significance of covariate correlation on both, as explained when discussing our approach above. A common set of variables for the models were then selected based on whether they were selected in either of the forward or backward stepwise regressions.

3. **Increasing the set of covariates with polynomial and interaction terms using algorithmic selection**

In a third step, the same method of stepwise regressions (backwards and forwards) was employed to augment the set of covariates by quadratic terms or interactions of variables that had already been selected in steps one and two. The rationale behind this is the fact that balance might only be achieved if the propensity score is estimated using non-linear transformations of the variables selected in the first two steps (Imbens and Rubin 2015, p. 287). Again, the stepwise regression approach helped to decide which of these non-linear terms were significant predictors of differences across control and treatment groups, and should therefore be controlled for.

The result of this process was the identification of a selection model comprising a set of covariates that were included in the first-stage estimation of the propensity score. This approach was conducted for every estimation strategy for each of the treatment effects, i.e. ATT1, ATT2, and ATT3. The full set of covariates finally included can be made available upon request. It is important to note, however, that good balancing properties using PSM also depend on the matching algorithm used in the second stage of the PSM analysis described in the next section.

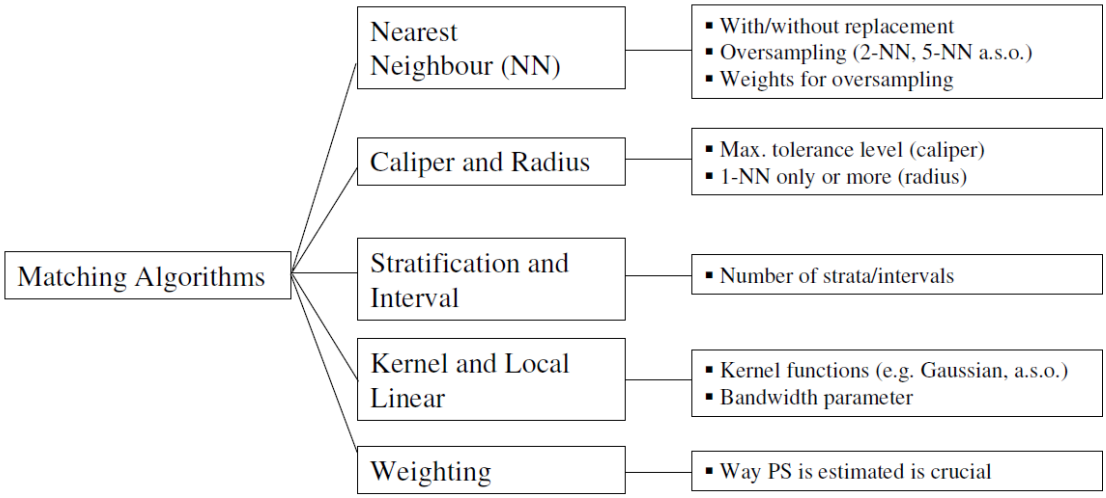
Second-stage algorithm selection

There are a variety of algorithms available to implement the second stage of PSM, i.e. to match control and treatment units to each other based on the propensity score estimated in the first stage. Figure 13 shows algorithm options and sub-options for each of these possibilities. It is beyond the scope of this report to explain in detail the technicalities of each of these approaches.[4] For all approaches the goal is to find appropriate, i.e. sufficiently similar, control group members for treatment group members. Differences between these approaches can be defined along three main dimensions: first, which estimated propensity scores are considered to be valid for inclusion in the analysis? Second, what is the appropriate range of propensity scores that define control comparators for treatment units? Finally, how are these comparators used when estimating the treatment effects?

The first dimension relates to the fact that within both control and treatment groups there could be estimated propensity scores that lie either at the upper or lower bound of the distribution, i.e. close to 0 or 1. For such values, there might not be an appropriately similar propensity score in the respective comparison group. However, for matching to work appropriately, there must be comparable propensity scores in both control and treatment groups – the so-called common support condition. Hence, matching algorithms employ cut-offs or trimming procedures by which some proportion of observations with propensity scores that are not comparable are dropped from the analysis.

[4] See Caliendo and Kopeinig (2005) for a summary overview.

Figure 13 Matching algorithms selection



NN: Nearest Neighbour, PS: Propensity Score

The second dimension relates to how units in the control group with propensity scores close to a treatment group observation are treated. For instance, kernel matching, as used in the main impact estimation for the PSM model, is a non-parametric matching estimator that uses the weighted averages of all individuals in the control groups to create the counterfactual outcome. The weights are determined by the distance between each individual from the control group and the participant observation for which the counterfactual is estimated. Therefore, higher weights are given to persons that are closer in terms of the propensity score of a treated individual (Caliendo and Kopeinig (2005), pp. 10–11). Alternatively, nearest neighbour (NN) matching with just one unit looks for the one control observation that has the closest propensity score to a treatment unit and compares the outcome measure for those observations. NN matching with more than one neighbour looks for several control units with similar propensity scores and compares the treatment outcome to an average of these neighbours. Caliper matching is similar to NN matching but does not include a fixed number of neighbours. Instead, the comparators are selected based on the maximum difference in propensity scores allowed.

Finally, the third dimension refers to how, once comparator units are found, the outcome measures are compared across treatment and control groups. For example, with NN matching and more than one neighbour simple averages are calculated. Similarly, with kernel functions a form of weighted averages are calculated to estimate treatment effects.

Selecting the appropriate matching algorithm for a PSM exercise is not straightforward and requires careful analysis of how well-balanced samples are after employing algorithms with certain sub-specifications. In general, however, the selection of models in this study was based on the fact that discriminating between models imposes a bias/variance trade-off in regard to the estimated treatment effect. For instance, in the extreme case of NN matching with just one neighbour, it could be that the NN is actually quite far away in terms of propensity scores, and hence a bad match. If this happens often, this could introduce bias into the estimation procedure. A solution to this could be to implement matching using several comparators in a caliper matching setting. However, this could decrease the number of available matches, which could increase the variance of the treatment estimate.

Kernel matching with appropriate trimming and enforcement of common support is a good compromise between these different approaches and was therefore selected as the main matching algorithm here. In order to find the optimal estimation model this study used different kernel matching algorithms with different bandwidths and trimming levels. These different results were then compared with respect to the best balancing properties, with the best performing approach being selected as the optimal.

Key PSM assumptions: common support and conditional independence

There are two key assumptions that need to hold for PSM to be a valid approach to estimating treatment effects: the common support assumption and the conditional independence assumption.

The **common support assumption** states that the estimated propensity score for all individuals in the treatment and control groups must lie within 0 and 1. Expressed differently, individuals in both groups must have a positive non-zero probability of belonging to either the treatment or control group, and the distribution of those probabilities across the two groups must be such that comparable individuals across the groups can be found. This can easily be enforced by only comparing observations with appropriate propensity scores.

The second key assumption is the **conditional independence assumption**, which posits that, once observable characteristics have been accounted for, the outcome measure is no longer related to the treatment status, except via the effect of the programme. In essence, this assumption states that once observable characteristics are appropriately controlled for, treatment status can be treated as if it was assigned randomly. As described above, PSM deals with this problem by comparing outcome measures across treatment and control groups only for individuals who are similar, i.e. by controlling for the important characteristics that are related to both treatment status and the outcome measure. The conditional independence assumption states that all important characteristics have been taken care of. This means that any bias that arises due to participation in the programme has been dealt with. Note that this includes biases that arise due to unobservable factors – PSM cannot control for these and the assumption is that once observable characteristics have been dealt with no unobservable bias remains.

The validity of any PSM approach therefore crucially depends on how well the approach reduces any imbalance between treatment and control groups. Under conditional independence – i.e. independence of the treatment assignment from outcome measures when controlling for covariates – the propensity score is a valid balancing score. Conditioning on this score appropriately means that bias will be removed between control and treatment groups. Hence, treatment and control groups will be balanced, i.e. they will have similar covariate distributions. This means that, across a variety of different characteristics, the treatment and control groups will be similar to each other.

Assessing the balance of covariates after matching is therefore a key step for any PSM analysis. The more balanced samples are after matching, the more plausible it is that the conditional independence assumption holds. As described above, however, balance also depends on the models and algorithms used to implement matching. The following paragraphs explain in detail how balance assessments were implemented and used in the current study.

Assessing balance

To select between different matching algorithms and to assess covariate balance after matching, this study compared matching models along a variety of dimensions. First, individual covariate balance was assessed across samples by looking at the standardised difference in means across treatment and control groups both before and after matching. This standardised difference is the difference in group averages over the square root of the average of the sample variances. If

samples are balanced, this difference should be small and matching should reduce this standardised difference as compared to the unmatched samples.

In addition, this study performed t-tests to assess whether differences across treatment and control groups were statistically significant. If balance is achieved with PSM, differences between treatment and control groups should be negligible and therefore should not be significantly different from zero.

In this context, the variance ratios of covariates of treated over control measures was also assessed. If there is perfect balance across samples, then covariates should be distributed equally and hence this ratio should be equal to one.

All of these measures give an indication of whether specific individual covariates are balanced across treatment and control groups. To assess overall variance, this study used two statistics that summarise covariate balance in the sample at hand: Rubin's B and Rubin's R. Rubin's B reflects the absolute standardised difference of the means of the propensity score in the treated and control groups (unmatched and matched). Rubin's R is the ratio of the treated to control variances of the propensity scores. Rubin (2001) suggests that the value of B should lie below 25 and that R should lie between .5 and 2 for overall balance to be sufficient. Together, Rubin's B and Rubin's R provide an informative indication of the trade-off between bias and variance across the treatment and control groups, as they change before and after the matching procedure. However, individual-level balance should always be assessed as the overall balance is only an approximation of goodness of fit.

Matching procedures were implemented using the psmatch2 package in Stata (Version 14.1) and balancing tests were carried out using the pstest package, which provides the results for all the statistics mentioned above.[5]

Finally, the distribution of propensity scores was also analysed graphically. Ideally, propensity scores should be distributed equally across treatment and control groups. Very skewed/diverging distributions could be an indication that balance has not been achieved successfully. The visual distribution of propensity scores was therefore taken into account in selecting the preferred estimation model for the impact analysis.

The results of balancing assessments are presented in Section 3.4.4 of this report.

What treatment effect does PSM estimate?

It is important to emphasise that the PSM approach used in this study works by looking for control units that can be compared to treatment units, and not the other way around. This means that it is assumed that treatment units are a given and control units need to be identified. Through finding matches for the treatment units in the pool of control units, the resulting estimate of the treatment effect is therefore the ATT. Depending on the exact definition of control and treatment units, this analysis produced three separate ATT estimates (see Section 3.4.1).

Considerations for applying PSM in the context of this evaluation

The following paragraphs further outline the approach to employing PSM in the context of the present evaluation. In particular, it is important to emphasise that we use PSM here to supplement our main impact analysis, which uses a DID approach, with additional insights relating to the relationship between IYCF counselling and exclusive breastfeeding among children aged 0–5

[5] See http://fmwww.bc.edu/repec/bocode/p/pstest.html for details.

months at endline. This means that we are running a PSM analysis on a limited subsample of our overall data, employing information from endline only.

In addition, it is important to mention that this means that we are implementing PSM on a sample of limited size, given that we have relatively few children aged 0–5 months in our sample. Note that the sample sizes are further limited by the different definitions of treatment effects. For ATT1, we have to limit our analysis to children whose mothers have received IYCF counselling in WINNN-supported LGAs and to children whose mothers have not received IYCF counselling in non-supported LGAs. Similarly, for ATT2, we have to limit our sample to children in WINNN-supported LGAs. For ATT3, we have to limit our sample to children whose mothers have not received IYCF counselling. Small sample sizes, in turn, have implications for how well PSM performs, and generally lead to a situation where achieving balance is difficult.

3.4.3 Caveats

Three key caveats related to the present estimation strategy need to be mentioned here. First, PSM only controls for observable characteristics that cause selection bias. This is a fundamental problem for any impact identification strategy that relies on controlling for observable variables. PSM helps to address this by allowing for extensive balancing checks after matching, which can provide substantial evidence for the fact that balance is achieved across a large variety of characteristics and – by implication – is likely to also extent to unobservables. In this study, such balancing checks were implemented. The results are presented below.

Second, as mentioned above, we are implementing PSM on a limited set of observations. PSM generally performs best with large samples, because it is then easier to find good comparisons between treatment and control groups, and hence we expect this to be slightly difficult in the current context. We implement balancing checks to ensure that comparisons are still acceptable for the purposes of this analysis, despite the limited sample size. All balancing test results are presented below.

Finally, calculating standard errors of estimated treatment effects using PSM methods is not straightforward. As Caliendo and Kopeinig (2005, p. 18) put it: ‘The problem is that the estimated variance of the treatment effect should also include the variance due to the estimation of the propensity score, the imputation of the common support, and possibly also the order in which treated individuals are matched’. These estimations increase the variation of the treatment effect estimates over and above normal sampling variation. In the literature, there is no consensus on how to take this into account.

A popular approach to solve this problem is to bootstrap standard errors for the estimated treatment effect (see Lechner, 2002). Each bootstrap draw re-estimates both the first and second stages of the estimation. This produces N bootstrap samples for which the ATT is estimated. The distribution of these means approximates the true sampling distribution, and therefore the standard errors of the population mean (Caliendo and Kopeinig, 2005, p. 18). This study followed this approach and implemented bootstrapping, using 200 repetitions, to estimate the standard errors of the estimated treatment effects. Note that, for the sake of completeness, this report shows both the bootstrapped and the non-bootstrapped standard errors below.

It is also important to note that there is no clear direction in which estimated standard errors should change due to bootstrapping. On the one hand, the additional variation taken into account should increase standard errors. On the other, bootstrapping generally makes estimates more precise, which tends to decrease standard errors. Overall, the direction of the change is not uniform. In fact,

the results show that, with bootstrapping, standard errors in some instances are smaller, and in some instances they are larger, than without bootstrapping.

3.4.4 Results

As an overview, Table 32 below again summarises how treatment and comparison groups are defined for the three different treatment estimates we produce in the context of the present analysis. Note, again, that we limit the analysis to endline only, where we expect effects of WINNN support to have materialised.

Table 32 Treatment and control group definitions for the different estimations – endline only

Treatment effect	Treatment group	Comparison group
$\mathbf{ATT1} = \mathbf{E}(\mathbf{Y}_{[\mathbf{IYCF}=\mathbf{1},\mathbf{T}=\mathbf{1}]} - \mathbf{Y}_{[\mathbf{IYCF}=\mathbf{0},\mathbf{T}=\mathbf{0}]})$	Children whose mothers **ever** received IYCF counselling in WINNN-**supported** LGAs	Children whose mothers **never** received IYCF counselling **non-supported** LGAs
$\mathbf{ATT2} = \mathbf{E}(\mathbf{Y}_{[\mathbf{IYCF}=\mathbf{1},\mathbf{T}=\mathbf{1}]} - \mathbf{Y}_{[\mathbf{IYCF}=\mathbf{0},\mathbf{T}=\mathbf{1}]})$	Children whose mothers **ever** received IYCF counselling in WINNN-**supported** LGAs	Children whose mothers **never** received IYCF counselling in WINNN-**supported** LGAs
$\mathbf{ATT3} = \mathbf{E}(\mathbf{Y}_{[\mathbf{IYCF}=\mathbf{0},\mathbf{T}=\mathbf{1}]} - \mathbf{Y}_{[\mathbf{IYCF}=\mathbf{0},\mathbf{T}=\mathbf{0}]})$	Children whose mothers **never** received IYCF counselling in WINNN-**supported** LGAs	Children whose mothers **never** received IYCF counselling **non-supported** LGAs

In what follows, we will be presenting the key results and balancing diagnostics for our preferred impact estimation strategies for each of ATT1, ATT2, and ATT3. In Volume 1, we present these results as a summary.

ATT1

For each treatment effect estimate, two sets of results are presented in this volume: first, the second stage and balancing results of our main strategy that is used in Volume 1, and, second, equivalent results of a robustness check strategy that shows how robust our results are to some changes in the estimation strategy.

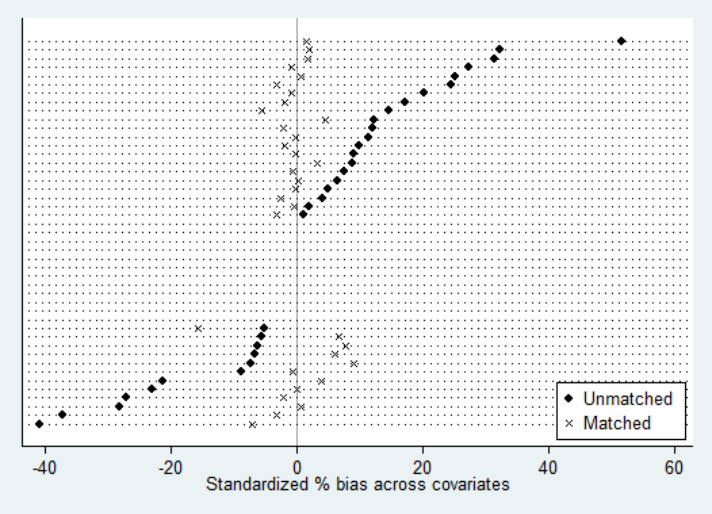
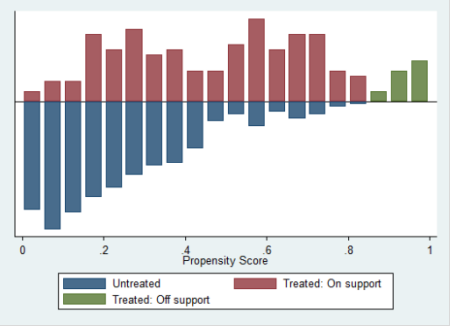
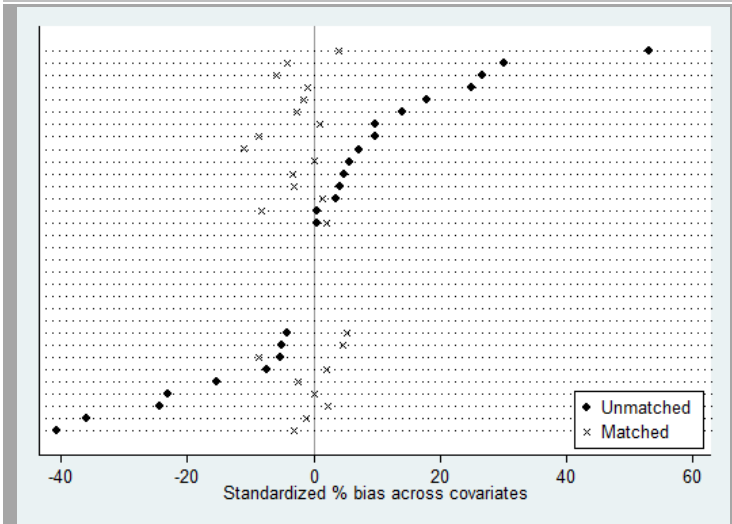
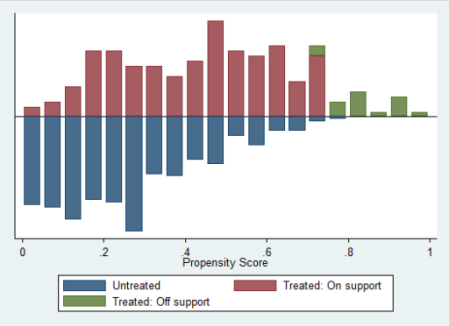
The second stage and balancing results for the main strategy and robustness check are presented as illustrated in Figure 14 for ATT1. The figure is divided into two panels: the top panel and the bottom panel show the main and robustness results, respectively. In this case, the robustness results are from an ATT estimation that changes the covariates controlled for to exclude interaction terms that might have been selected in the first stage. The format for each panel is as follows:

- The first graph on the left-hand side indicates how individual variables balance before and after matching. The x-axis displays the standardised bias, which is the percentage difference of the sample means in the treated and non-treated (unmatched or matched) subsamples as a percentage of the square root of the average of the sample variances in the treated and non-treated groups (Rosenbaum and Rubin, 1985). In Figure 14 below, for example, the unmatched samples display large imbalances, with standardised bias being present across many of the covariates of interest. However, once matching takes place, the standardised imbalances are reduced significantly.
- The second graph, on the right-hand side, shows the distribution of propensity scores across treatment and control groups. This graph visually confirms that, after dropping observations that are off common support, both treatment and control groups contain observations with

propensity scores that are similar, which is an indication of overall balance. Although the distributions of propensity scores across treatment and control groups would ideally be symmetric, the presence of some level of skewness does not put at risk the estimation procedure, as indicated by the balance achieved for each covariate and the overall values of Rubin's R and B after matching.

- The remaining rows on the right-hand side display information related to the PSM model. The bandwidth and level of trimming for the optimal PSM model can be found in the first two rows. For example, the optimal model has a bandwidth of 6 and a trimming value of 10 for the baseline sample in Figure 14. This relates to the fact that we are using kernel matching in our estimation procedure. This is then followed by the number of observations on common support in the next row, and then the Rubin's R and Rubin's B values both before and after matching. Generally, a Rubin's B score under 25 after matching is desirable, while a Rubin's R score between 1 and 1.25 is the preferred range after matching (Rubin, 2001). The unmatched samples are particularly unbalanced: for instance, the Rubin's B score for the main strategy is 107.44. However, the Rubin's B score after matching is 29.8, which shows how matching removes a significant amount of the previous imbalances. Note here that Rubin's B is not, however, under 25, which indicates that some imbalance remains. In the robustness analysis, the score moves to under 25, indicating somewhat better balance. Given that our ATT estimate here does not vary significantly, we conclude that despite remaining imbalances our estimate is robust.
- Finally, the remaining rows on the left-hand side of each panel indicate the ATT for each corresponding estimation and the associated standard errors. Given that it is not definitively clear how to produce standard errors for PSM, both bootstrapped and non-bootstrapped standard errors are presented for robustness purposes.

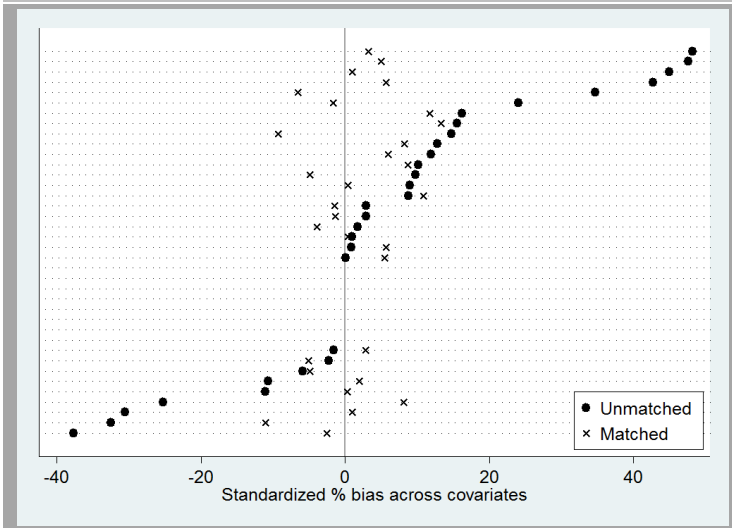
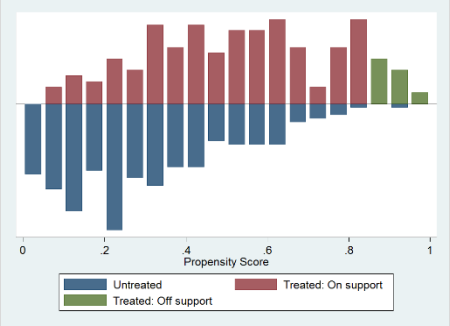
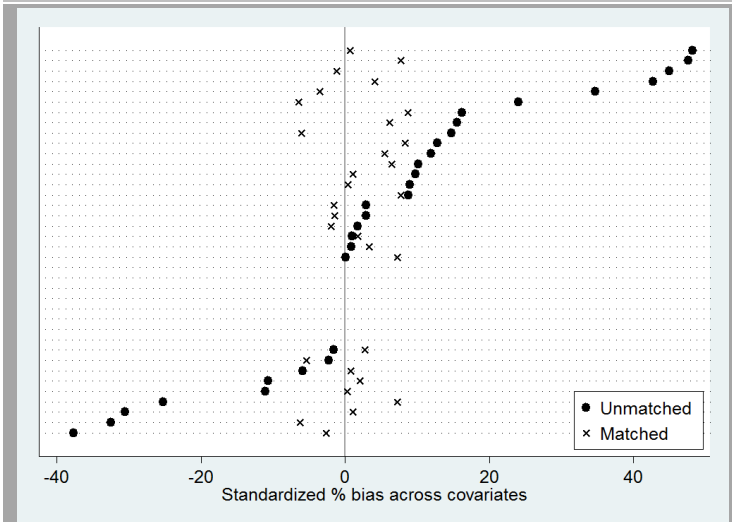
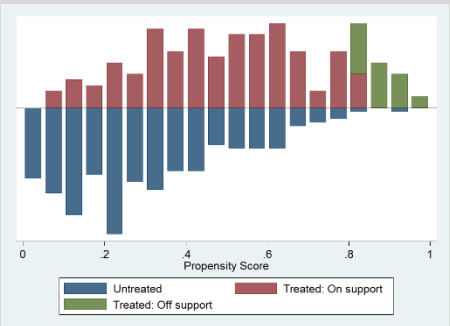
Figure 14 ATT1 estimation (main estimate of interest) on exclusive breastfeeding

Standardised bias across covariates and ATT		Balancing		
Main strategy				
				
		Bandwidth		4
		Trimming		10
		N on common support		509
		Rubin's B Rubin's R	[before matching]	107.44 1.7
ATT **Standard error (bootstrapping)** **Standard error (no bootstrapping)**	**19.35** **(3.3)** **(4.1)**	**Rubin's B** **Rubin's R**	**[after matching]**	**29.88** **0.91**
Robustness check – estimation without interaction terms				
				
		Bandwidth		4
		Trimming		10
		N on common support		512
		Rubin's B Rubin's R	[before matching]	103.98 1.07
ATT **Standard error (bootstrapping)** **Standard error (no bootstrapping)**	**19.00** **(4.1)** **(3.9)**	**Rubin's B** **Rubin's R**	**[after matching]**	**21.04** **1.13**

ATT2

The figure below shows results from the estimation procedure for ATT2. Note that this is comparing children whose mothers received IYCF counselling with mothers who did not receive IYCF counselling within WINNN-supported LGAs. Aside from the fact that we find a significant positive effect on exclusive breastfeeding, the following point needs to be emphasised: balancing results indicate that after matching samples are not perfectly balanced. In fact, as seen in Figure 15, in both cases, Rubin's B is larger than 25, although the robustness check strategy is very close to this threshold. Hence, imbalances remain even after matching. The fact that we find very similar and significant treatment effects in both strategies, however, gives us confidence in our main conclusion that within WINNN-supported LGAs IYCF counselling attendance makes a positive difference in terms of the propensity of mothers to exclusively breastfeed.

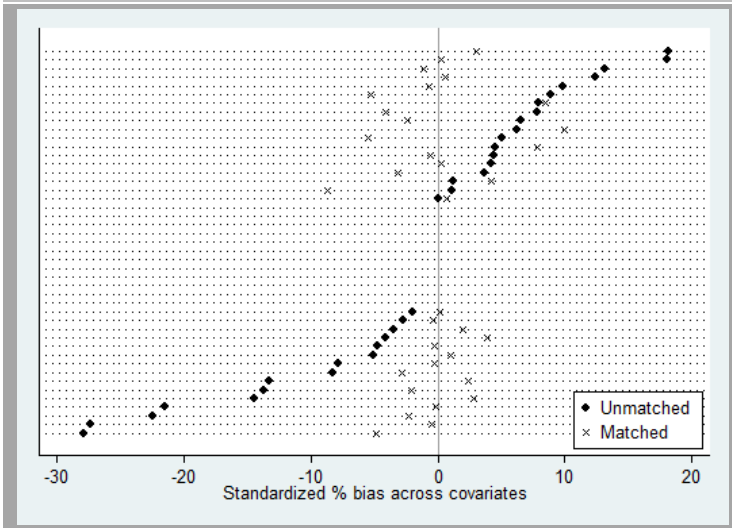
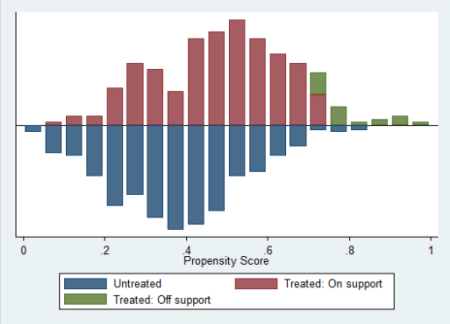
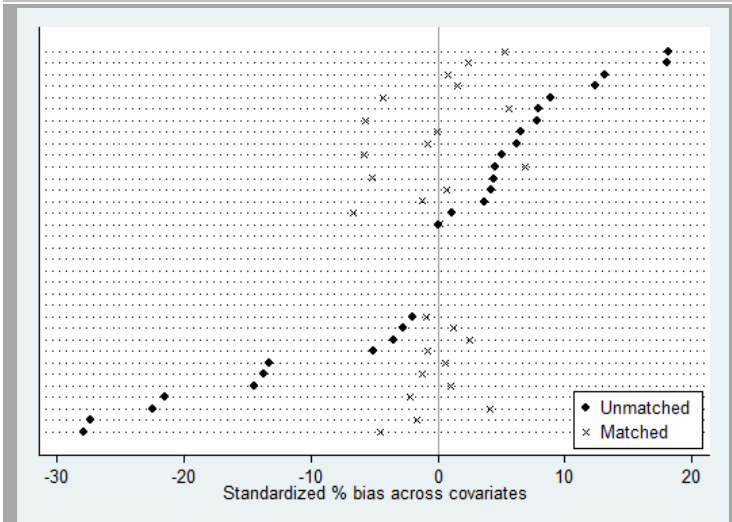
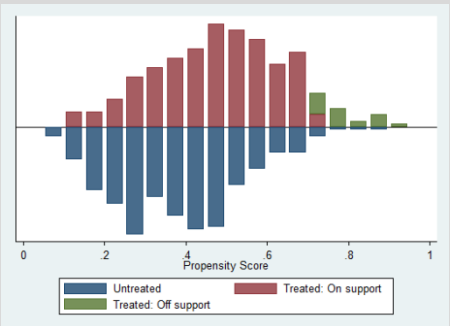
Figure 15 ATT2 estimation (estimating effects of the IYCF interventions within WINNN-supported LGAs) on exclusive breastfeeding

Standardised bias across covariates and ATT		Balancing		
Main strategy				
				
		Bandwidth		4
		Trimming		10
		N on common support		408
		Rubin's B Rubin's R	[before matching]	111.84 1.1
ATT **Standard error (bootstrapping)** **Standard error (no bootstrapping)**	**12.0** **(4.4)** **(4.6)**	**Rubin's B** **Rubin's R**	**[after matching]**	**33.46** **1.2**
Robustness check – trimming at 15				
				
		Bandwidth		4
		Trimming		15
		N on common support		399
		Rubin's B Rubin's R	[before matching]	111.98 1.1
ATT **Standard error (bootstrapping)** **Standard error (no bootstrapping)**	**10.42** **(4.5)** **(4.4)**	**Rubin's B** **Rubin's R**	**[after matching]**	**26.04** **1.49**

ATT3

The figure below presents equivalent results for the third impact estimation strategy to estimate ATT3. Note that this involves comparing individuals who did not receive IYCF counselling across supported and non-supported LGAs. The results show that balancing is achieved both for the main and the robustness strategy – which in this case involves excluding interaction terms. In fact, balancing diagnostics show that this strategy is the most balanced after matching. This means that individuals who are included in this analysis are similar to each other.

Figure 16 ATT3 estimation (estimating WINNN effects across LGAs) on exclusive breastfeeding

Standardised bias across covariates and ATT		Balancing		
Main strategy				
				
		Bandwidth		4
		Trimming		8
		N on common support		592
		Rubin's B Rubin's R	[before matching]	69.42 0.91
ATT **Standard error (bootstrapping)** **Standard error (no bootstrapping)**	**6.5** **(2.6)** **(2.6)**	**Rubin's B** **Rubin's R**	**[after matching]**	**23.28** **1.36**
Robustness check – no interaction terms				
				
		Bandwidth		4
		Trimming		8
		N on common support		592
		Rubin's B Rubin's R	[before matching]	71.74 1.09
ATT **Standard error (bootstrapping)** **Standard error (no bootstrapping)**	**6.00** **(2.6)** **(2.5)**	**Rubin's B** **Rubin's R**	**[after matching]**	**18.12** **1.11**

3.5 Supplementary correlation analysis

The purpose of this section is to further explain the results of the correlation analysis that are presented in Box 2 of Volume 1.

3.5.1 Aim

Given the impact results presented in Volume 1 and above in this volume, the purpose of this additional correlation analysis is to investigate what characteristics of mothers and children living in WINNN-supported LGAs are correlated with different levels of key impact indicators. Note that this supplementary analysis was **implemented in response to explicit requests by stakeholders during this evaluation** and that the key objective is to identify the characteristics of individuals in WINNN-supported LGAs that might explain some of the variation in key indicators in those areas at endline.

This analysis is carried out along three dimensions:

- First, awareness among mothers of WINNN-supported services.
- Second, uptake of WINNN-supported services.
- Third, provision of key nutritional care practices by mothers.

However, instead of looking at certain background characteristics individually, the objective of this section is to identify variables that are significantly correlated with impact and outcome indicators while holding others constant, i.e. to analyse this in a multivariate context.

It is important to mention that this is not a causal analysis, i.e. that variables identified here are not necessarily causally influencing outcome or impact indicators. As the title of this section suggests, this analysis identifies background characteristics which are, statistically speaking, significantly related to the likelihood of being aware of WINNN-supported services, of taking such services up, or of providing key nutritional care practices, holding all other variables in the model constant. Background characteristics identified in such a manner can shed further light on the context of the implementation of WINNN, and potential factors that are related to impact and outcome indicators of interest, and could therefore provide insights that are relevant for future programming.

3.5.2 Methodology

In order to identify such background characteristics, we limit this analysis to data from WINNN-supported LGAs collected at endline. This is because individuals from non-supported LGAs are, in theory, not exposed to WINNN services and because, at baseline, the correlations of interest will not have materialised yet.

We also limit this supplementary analysis to a small set of key outcome indicators, which are presented in

Table 33 below. As mentioned above, we consider these to be proxies for mothers' awareness of WINNN-supported services, mothers' uptake of, or exposure to, WINNN-supported services, and mothers' provision of positive nutritional practices.

Table 33 Outcome indicators considered in this analysis

Outcome indicators
Awareness:
Proportion of mothers who have ever heard about food demonstration sessions in the community
Proportion of mothers who have ever heard about MNCHW events
Exposure and uptake:
Proportion of mothers who have ever attended IYCF counselling in the community
Proportion of children who have ever had their MUAC measured
Proportion of mothers who have attended MNCHW events
Good nutritional care:
Exclusive breastfeeding (children 0–5 months)
Early initiation (< 1 hour) to breastfeeding (children 0–23 months)

In order to identify key background characteristics which are correlated with the above outcome indicators, we use a simple regression framework, which can be defined as follows:

$$(6)\quad Y = X\beta + \epsilon,$$

where Y is the outcome variable of interest, X is a vector of relevant covariates that we want to examine, and ϵ is the error term. Our aim is to estimate β using OLS and to identify which of the coefficients are significantly different from zero.

Of course, doing this requires selecting the set of covariates in X that are used to estimate this equation. In order to prevent the identification of spurious relationships we run a set of five multivariate regression model specifications, i.e. with five different compositions of X, for each of the outcome indicators listed in the table above:

- First, a simple OLS regression that includes variables in X that could, from a theoretical perspective, be of relevance. This includes information such as, for examples the state where mothers and children live, the mother's age, the education level of mothers, the household's wealth quintile, and the distance from the nearest WINNN-supported health facility. By doing this, we include variables that have previously been used to disaggregate outcome indicators in Volume 1 of this report.
- In addition, two regression models that include these theoretical variables, plus variables that were automatically selected by backward and forward stepwise regressions from a full set of over 100 possible variables. The vector of covariates is hence a composite of theoretically and algorithmically selected variables. Note that we deal with multicollinearity here by excluding highly collinear variables from this approach by looking at the variance inflation factor of variables within each regression.
- A regression model that includes the set of theoretical variables plus variables that were selected by a LASSO regression from a full set of over 100 possible variables. Again, the vector of covariates here is composed of variables selected based on theoretical priors and the LASSO estimation procedure.
- Finally, a regression model in which the full set of covariates was selected by a LASSO regression, without forcing theoretically defined variables into the model.

Note that all regressions are implemented using the survey settings option in Stata (Version 14.1), i.e. taking into account the survey structure of our data.

In order to prevent spurious associations being picked up, we consider variables to be significantly correlated with the outcome indicator for the purposes of Volume 1 only if such significant

relationships are picked up by three or more of the specifications mentioned above. The full results from these different specifications for each of the outcome indicators are listed in Section 3.5.4.

3.5.3 Caveats

Two key caveats need to be borne in mind for the purposes of this analysis:

- First, it is important to reiterate that this analysis does not claim to identify causal effects. This means that when we say, for example, that having primary or secondary education is positively related to the proportion of mothers having heard about food demonstrations, this does not mean that higher education 'caused' mothers to be more likely to hear about food demonstrations. It is possible that higher education levels are related to some other variable not included in the model (for example, because it is unobservable – such as the motivation of mothers to learn about new things), which in turn is also related to the likelihood of being aware of food demonstrations. Making causal claims would require much more extensive modelling, which was outside the scope of this supplementary analysis, and, potentially, an impact evaluation design.
- Second, there is the possibility that the algorithmic approach designed above will pick up spurious relationships between the background characteristics of individuals and the outcome variables. Spurious relationships are defined as relationships that are identified as being significant just because of a feature of the algorithm used or the sampled data, but that do not represent an actual, true, relationship in the underlying population. In a classical regression context, endogeneity can lead to such identification problems, which includes omitted variable bias, simultaneity and measurement error. Stepwise regression approaches are prone to this problem. Because of this, we adopt a conservative approach in this analysis: first, we complement stepwise regressions with LASSO regressions, which are not prone to this problem; second, we only characterise relationships as being significant and report them in Volume 1 when variables are significant across three regression specifications or more. Significance here means p-values of under 0.05, i.e. significance levels of 5%.

3.5.4 Results

This section presents the results from OLS regressions for each of the indicators listed in

Table 33. The tables should be interpreted as follows:

- The first column on the left lists all variables that were fed into the different OLS regressions (1) to (5). The variables in bold are the ones that we identified as being relevant, from a theoretical perspective. Note that we purposefully kept this set of variables limited in terms of the number of variables included.
- Each of the columns (1) to (5) lists the coefficients estimated in each of the regressions, including significance stars when coefficients were significant at the 5% (*), 1% (**), or 0.1% (***) level. Column (1) presents results from an OLS regression that only includes variables for which we have a prior theoretical basis. Column (2) presents results from a regression that includes these theoretical variables plus a set of variables that were identified to be relevant in a backward stepwise regression algorithm that was run on a larger set of over 100 covariates.[6] Column (3) presents results from a regression that includes the 'theory' set plus variables that were selected in a forward stepwise regression algorithm. Column (4) presents results from a regression model that includes variables selected by a LASSO regression plus the theoretical variables. Finally, column (5) presents results from a regression in which no theoretical prior was set: instead, LASSO was allowed to run on a full set of covariates, i.e. no variables were forced into the regression.
- For each of the regressions we also report R-squared and the number of observations used in the last two rows.
- Background characteristics that we consider to be of relevance are highlighted in blue in the tables. For each of these, this means that, even when holding all other variables in the models below constant, these background characteristics are still statistically significantly correlated with the outcome variable we are looking at. The sign of the coefficients tells us whether that relationship is positive or negative.

Table 34 Correlation analysis on the proportion of mothers who have ever heard about food demonstration sessions in their community

Background characteristic of mother	(1)	(2)	(3)	(4)	(5)
	Theoretical model	Backward stepwise	Forward stepwise	LASSO + theoretical	LASSO alone
Katsina	0.228***	0.177***	0.170***	0.165***	0.181***
Kebbi	-0.061	-0.0554	-0.0664	-0.0615	-0.0786*
Zamfara	-0.0296	-0.052	-0.0576	-0.0722	-0.0742
Aged 20 to 24 years	0.0338	0.0376	0.0392	0.0112	
Aged 25 to 34 years	0.0585	0.0663	0.0685	0.0249	
Aged more than 35 years	0.0438	0.0333	0.0344	-0.00107	
Primary schooling	0.116**	0.108*	0.106*	0.109*	
Secondary schooling	0.122*	0.130*	0.122*	0.0914	
Distance to nearest WINNN facility in kilometres (km)	0.00117	0.000303	0.000449	0.000397	
Second wealth quintile	0.0763*	0.0473	0.0496	0.0574	
Third wealth quintile	0.0802*	0.0819*	0.0764	0.0633	
Fourth wealth quintile	0.0651	0.0286	0.0291	0.0487	
Fifth wealth quintile	0.0384	-0.0142	-0.0116	0.015	
Number of children	0.00922	0.00411	0.00488	0.00632	
Polygamous marriage		0.0988***	0.102***	0.105***	0.105***

[6] The list can be provided upon request.

Household head with any formal education		0.0624	0.0598	0.0494	0.0623*
Simple violence index, violence sometimes OK		0.105***			
Travel time to nearest malaria facility 1 to less than 2 hours		-0.0668	-0.0622		
It is OK to beat wife if she burns the food		0.104**			
Household owns cow		0.0721**	0.0677*		
Mother with Islamia education		-0.0877***	-0.0905***	-0.0746**	-0.0943***
Travel time to nearest malaria facility 5 hours or more		-0.191***	-0.118*		
Household with 3 to 4 bedrooms		-0.00754	-0.00661		
Moderate household hunger (FAO hunger scale)		-0.155*			
Did anybody ever go to sleep hungry in last four weeks?		0.164			
It is OK to beat wife if she neglects children			0.0947***		
Household owns horse				-0.0483	
Mother is widowed					-0.129
Constant	0.051	0.0148	0.0416	0.0975	0.194***
R-squared	0.104	0.15	0.139	0.126	0.114
N	2297	2109	2109	2168	2197

Note: Results from OLS regression on ORIE endline data collected within WINNN-supported LGAs. Surveysettings taken into account, including clustering of standard errors. Significance stars defined as follows: * = p<0.05; ** p<0.01; *** p<0.001.

Table 35 Correlation analysis on the proportion of mothers who have ever heard about MNCHW events

Background characteristic of mother	(1)	(2)	(3)	(4)	(5)
	Theoretical model	Backward stepwise	Forward stepwise	LASSO + theoretical	LASSO alone
Katsina	0.0625	0.00242	0.00459	0.0215	
Kebbi	-0.194***	-0.182***	-0.184***	-0.190***	-0.206***
Zamfara	-0.036	-0.0342	-0.0471	-0.0814	
Aged 20 to 24 years	-0.0343	-0.0227	-0.0418	-0.0593	
Aged 25 to 34 years	0.0381	0.0352	0.0227	0.0112	
Aged more than 35 years	-0.0221	-0.0681	-0.0872	-0.0586	
Primary schooling	0.110**	0.110*	0.0934*	0.0931*	
Secondary schooling	0.131	0.126	0.11	0.0967	
Distance to nearest WINNN facility in km	0.000782	0.00262	0.00288	0.00103	
Second wealth quintile	0.113**	0.0921*	0.0927*	0.0861*	
Third wealth quintile	0.121**	0.115*	0.109*	0.0885	
Fourth wealth quintile	0.176**	0.158**	0.152*	0.149*	
Fifth wealth quintile	0.182***	0.154*	0.151*	0.144**	
Number of children	0.0129	0.0129	0.0133	0.0123	
Polygamous marriage		0.0772**	0.0807**	0.0861**	
Household owns a car		0.0666	0.0577		

Travel time to nearest malaria facility 5 hours or more		-0.526***	-0.413***		
Travel time to nearest malaria facility 1 to less than 2 hours		-0.0699	-0.0591		
Household head with any formal education		0.0581	0.0674	0.0581	0.101**
Mother's main economic activity is farming/herding for subsistence		-0.328	-0.332		
Household head is divorced/separated		0.591***	0.492**		
Age of household head in years		0.00799***	0.00773***		
Household with three to four bedrooms		-0.0166**	-0.0169***		
Household head with Islamia education		0.0974**			
Household cooks indoors		-0.0619	-0.0619		
Household head is widowed		0.519***	0.509***		
Travel time to nearest malaria facility 30 to less than 60 minutes.		-0.0799	-0.0775		
Household has elderly household head (65 years or more)		-0.173**	-0.164**		
Mother with Islamia education			-0.0908**	-0.0978**	-0.0976**
Travel time in minutes from community to the nearest market			-0.0000544		
Travel time to nearest malaria facility less than 30 minutes					0.107**
Constant	0.258***	-0.107	0.0233	0.305***	0.379***
R-squared	0.082	0.131	0.13	0.101	0.07
N	2297	2024	2035	2191	2197

Note: Results from OLS regression on ORIE endline data collected within WINNN-supported LGAs. Survey settings taken into account, including clustering of standard errors. Significance stars defined as follows: * = p<0.05; ** p<0.01; *** p<0.001.

Table 36 Correlation analysis on the proportion of mothers who have ever attended IYCF counselling in the community

Background characteristic of mother	(1) Theoretical model	(2) Backward stepwise	(3) Forward stepwise	(4) LASSO + theoretical	(5) LASSO alone
Katsina	0.0727	0.0558	0.0871	0.0725	0.125*
Kebbi	-0.0626	-0.0294	-0.0149	-0.0289	
Zamfara	-0.125*	-0.100*	-0.0976	-0.112*	-0.0920*
Aged 20 to 24 years	0.0491	0.0673	0.0486	0.0299	
Aged 25 to 34 years	0.0434	0.0738	0.0512	0.0364	
Aged more than 35 years	0.12	0.154*	0.119	0.111	
Primary schooling	0.0755	0.0519	0.054	0.0673	
Secondary schooling	0.0429	0.013	0.0187	-0.00626	
Distance to nearest WINNN facility in km	0.0000809	-0.000971	-0.000125	-0.000449	
Second wealth quintile	0.0551	0.0259	0.0462	0.0335	
Third wealth quintile	0.104	0.106*	0.0805	0.076	
Fourth wealth quintile	0.0729	0.0539	0.0861	0.0655	
Fifth wealth quintile	0.0705	0.0485	0.0881	0.0497	

Number of children	0.00261	0.0013	0.00411	0.00156	
Did anybody in the household go to sleep hungry in the last four weeks?		0.139			
Travel time to nearest malaria facility 5 hours or more		-0.227**	-0.245***		
Travel time to nearest malaria facility 1 to less than 2 hours		-0.120*	-0.114*		
Household owns motorcycle/scooter		0.0840**			
Nearest market lies within 60 minutes of the community		-0.136***	-0.138**		
Household owns a boat/canoe		-0.250***	-0.0494		
Moderate household hunger (FAO hunger scale)		-0.137			
It is OK to beat wife if she burns the food		0.0881**	0.0892**		
Household with three to four bedrooms		-0.00734			
Mother with Islamia education		-0.0843**	-0.0877**	-0.0831**	
Household head with any formal education		0.0572	0.0546	0.0384	0.0779*
Household head's main economic activity is 'others'			-0.00605		
Household owns horse				-0.076	
Constant	0.258***	-0.107	0.0233	0.305***	0.379***
R-squared	0.082	0.131	0.13	0.101	0.07
N	2297	2024	2035	2191	2197

Note: Results from OLS regression on ORIE endline data collected within WINNN-supported LGAs. Survey settings taken into account, including clustering of standard errors. Significance stars defined as follows: * = p<0.05; ** p<0.01; *** p<0.001.

Table 37 Correlation analysis on the proportion of children who have ever had their MUAC measured

Background characteristic of mother	(1)	(2)	(3)	(4)	(5)
	Theoretical model	Backward stepwise	Forward stepwise	LASSO + theoretical	LASSO alone
Katsina	-0.00159	0.000313	0.000313	0.006	
Kebbi	-0.103***	-0.0826***	-0.0826***	-0.0962***	-0.101***
Zamfara	0.00579	0.0264	0.0264	0.0138	
Sex of child	-0.0192	-0.0212	-0.0212	-0.0227	
Mother's age in years	0.00550***	0.00441**	0.00441**	0.00463***	
Mother has primary schooling	0.0208	0.0119	0.0119	0.0262	
Mother has secondary schooling	-0.044	-0.0668	-0.0668	-0.0482	
Distance of household to nearest WINNN facility in km	0.00171	0.00194	0.00194	0.00184	
Second wealth quintile	0.007	-0.00525	-0.00525	0.00831	
Third wealth quintile	-0.0168	-0.0279	-0.0279	-0.0155	
Fourth wealth quintile	-0.018	-0.0453	-0.0453	-0.0222	
Fifth wealth quintile	-0.0356	-0.0690*	-0.0690*	-0.0415	
Mother's main economic activity is farming/herding for subsistence		-0.0838	-0.0838		
Mother's main economic activity is commercial farming/herding		-0.168***	-0.168***		
Travel time to nearest malaria facility 5 hours or more		-0.344***	-0.344***		

Mother thinks it is OK to beat wife if she goes out without telling husband		0.0253	0.0253		
Child's age in months		0.00713***	0.00713***	0.00709***	0.00723***
Child's mother went to antenatal care (ANC) sessions at least four times		0.0927***	0.0927***		
Mother is widowed				-0.143*	-0.113
Constant	0.057	-0.0645	-0.0645	-0.0348	0.0660***
R-squared	0.026	0.077	0.077	0.063	0.05
N	2614	2613	2613	2613	2625

Note: Results from OLS regression on ORIE endline data collected within WINNN-supported LGAs. Survey settings taken into account, including clustering of standard errors. Significance stars defined as follows: * = p<0.05; ** p<0.01; *** p<0.001.

Table 38 Correlation analysis on the proportion of mothers who have attended MNCHW events

Background characteristic of mother	(1)	(2)	(3)	(4)	(5)
	Theoretical model	Backward stepwise	Forward stepwise	LASSO + theoretical	LASSO alone
Katsina	0.0700*	0.0717*	0.0717*	0.0626	0.0847**
Kebbi	-0.0166	-0.0272	-0.0272	-0.0229	
Zamfara	-0.0328	-0.0351	-0.0351	-0.0362	
Aged 20 to 24 years	-0.0345	-0.0384	-0.0384	-0.0365	
Aged 25 to 34 years	-0.0408	-0.0436	-0.0436	-0.0422	
Aged more than 35 years	-0.0461	-0.057	-0.057	-0.0524	
Primary schooling	0.0571*	0.0497	0.0497	0.0522	
Secondary schooling	0.00501	-0.00308	-0.00308	-0.00218	
Distance to nearest WINNN facility in km	-0.000845	-0.000476	-0.000476	-0.000815	
Second wealth quintile	-0.000815	-0.00728	-0.00728	-0.00825	
Third wealth quintile	-0.00534	-0.0147	-0.0147	-0.0175	
Fourth wealth quintile	0.0116	-0.00154	-0.00154	-0.00102	
Fifth wealth quintile	-0.0506	-0.0647*	-0.0647*	-0.0651*	
Number of children	0.00644	0.00794	0.00794	0.00731	
Mother has heard of MNCHW events	0.286***	0.280***	0.280***	0.282***	0.280***
There was a flood in the community in the last 12 months		0.0342	0.0342		
Travel time to nearest malaria facility 1 to less than 2 hours		-0.0959***	-0.0959***	-0.0957***	-0.0509*
Mother with Islamia education		-0.0341*	-0.0341*	-0.0337*	-0.0227
Travel time to nearest malaria facility less than 30 minutes					0.0466**
Constant	0.0101	0.0223	0.0223	0.0433	-0.0393*
R-squared	0.221	0.234	0.234	0.231	0.223
N	2297	2295	2295	2295	2301

Note: Results from OLS regression on ORIE endline data collected within WINNN-supported LGAs. Survey settings taken into account, including clustering of standard errors. Significance stars defined as follows: * = p<0.05; ** p<0.01; *** p<0.001.

Table 39 Correlation analysis on exclusive breastfeeding among children aged 0–5 months

Background characteristic of mother	(1)	(2)	(3)	(4)	(5)
	Theoretical model	Backward stepwise	Forward stepwise	LASSO + theoretical	LASSO alone
Katsina	0.0327	0.019	0.0434	0.0395	
Kebbi	-0.0407	-0.011	-0.0317	-0.0321	
Zamfara	0.0719	0.105	0.0935	0.0876	
Sex of child	-0.0485	-0.0533	-0.0578	-0.0517	
Mother's age in years	0.0000498	0.000466	-0.00146	-0.00118	
Mother has primary schooling	0.0392	0.0249	0.0372	0.049	
Mother has secondary schooling	0.119	0.157	0.0826	0.137	
Distance of household to nearest WINNN facility in km	0.000955	0.000824	-0.000183	0.000984	
Second wealth quintile	0.0945	0.0843	0.068	0.0849	
Third wealth quintile	0.177**	0.173**	0.139*	0.159**	
Fourth wealth quintile	0.0804	0.0829	0.0388	0.0635	
Fifth wealth quintile	0.0329	0.076	-0.00243	0.0196	
Mother ever attended training on IYCF practices in community	0.0975	0.09	0.107*	0.0981	
Household owns air conditioner		-0.236*		-0.21	-0.172***
Household owns iron		-0.128*			
Child's mother went to ANC session at least four times		0.108			
Someone in the household went 24 hours without eating in the last four weeks			-0.139***	-0.161***	-0.196***
Household head is widowed			-0.311***		
Mother is younger than 18 years			-0.184*	-0.167*	-0.134*
Child received/mother bought iron supplements during pregnancy			0.104*		
Mother has to ask for permission to go to the village			0.171*		
Severe household hunger (FAO hunger scale)				0.0155	-0.0723
Constant	0.127	0.0927	-0.0241	0.18	0.209***
R-squared	0.06	0.08	0.091	0.073	0.015
N	448	448	426	448	451

Note: Results from OLS regression on ORIE endline data collected within WINNN-supported LGAs. Survey settings taken into account, including clustering of standard errors. Significance stars defined as follows: * = p<0.05; ** p<0.01; *** p<0.001.

Table 40 Correlation analysis on early initiation (<1 hour) to breastfeeding among children aged 0–23 months

Background characteristic of mother	(1)	(2)	(3)	(4)	(5)
	Theoretical model	Backward stepwise	Forward stepwise	LASSO + theoretical	LASSO alone
Katsina	0.278***	0.266***	0.266***	0.262***	0.224***
Kebbi	0.0438	0.0342	0.0342	0.0444	
Zamfara	0.142**	0.124*	0.124*	0.136**	
Sex of child	-0.0162	-0.0211	-0.0211	-0.0157	

Mother's age in years	0.00387*	0.00349	0.00349	0.00325	
Mother has primary schooling	-0.00695	-0.014	-0.014	-0.0437	
Mother has secondary schooling	0.0197	-0.0781	-0.0781	-0.111	
Distance of household to nearest WINNN facility in km	-0.00268	-0.000509	-0.000509	-0.00148	
Second wealth quintile	0.0534	0.0407	0.0407	0.0138	
Third wealth quintile	0.131**	0.138**	0.138**	0.112*	
Fourth wealth quintile	0.184***	0.182***	0.182***	0.148**	
Fifth wealth quintile	0.135*	0.0787	0.0787	0.053	
Mother ever attended training on IYCF practices in community	0.0628	0.0726	0.0726	0.0456	
Household owns air conditioner		0.103	0.103		
Travel time from community to the nearest market in minutes		-0.000846*	-0.000846*		
Community has experienced a drought in the last 12 months		0.00214	0.00214		
Household head is widowed		-0.0858	-0.0858	-0.148	-0.113
Travel time to nearest malaria facility is more than 5 hours		0.782***	0.782***		
Household owns a PC or laptop		0.333**	0.333**	0.312**	0.335***
Household owns a horse				0.0409	
Mother with Islamia education				-0.0653	
Marriage age of mother in years				0.0141	
Child's mother went to ANC sessions at least four times				0.0850*	
Constant	0.056	0.126	0.126	-0.116	0.309***
R-squared	0.089	0.107	0.107	0.109	0.058
N	1758	1699	1699	1675	1737

Note: Results from OLS regression on ORIE endline data collected within WINNN-supported LGAs. Survey settings taken into account, including clustering of standard errors. Significance stars defined as follows: * = p<0.05; ** p<0.01; *** p<0.001.

4 The impact evaluation results in context

4.1 How generalisable are the findings of this impact evaluation?

As mentioned in Section 3.1.1, this study's sample was drawn so as to provide representative estimates of WINNN programme LGAs at baseline in four states in northern Nigeria, and is therefore not representative of northern Nigeria overall, or of the four states included in the study. For these reasons, the external validity of the study, in a statistical sense, is limited: the findings are statistically representative for WINNN-supported LGAs and not beyond those areas.

However, generally speaking, it is still useful to consider the insights derived from this study within the broader geographical context of northern Nigeria and the Sahel region. This is because a number of states in northern Nigeria, and across the western Sahel, share similar characteristics in terms of seasons, health service provision, and cultural practices and beliefs, which might give an indication of what can be achieved should a programme similar to WINNN be scaled up or replicated elsewhere.

That said, before any inferences are transferred to another context directly from our findings, it is imperative to consider the context of how WINNN was implemented and its operational effectiveness. Findings from ORIE's qualitative workstream, for example, show that even within WINNN states, cultural differences might have played a role in how effective the implementation of WINNN-supported activities was across those states. Monitoring the effective implementation of the WINNN programme is the explicit remit of the Operations Research workstream within the ORIE project and findings will be shared separately.

In terms of the implications of this study's findings for Yobe, the fifth WINNN state, which remains outside the scope of the ORIE project, it is difficult to make any inferences as the continuing level of insecurity in Yobe makes this a very different operational context to the four WINNN states included in the present study.

4.2 Comparison with other surveys

This section cross-references key indicators from the endline study with other surveys in Nigeria, to provide further context to our analysis.

While comparisons are useful in triangulating findings and validating results, they must be interpreted carefully. There are four important considerations that must be taken into account when interpreting the cross-reference tables below, as follows:

1. The population base for a particular indicator may be different across surveys.

- For example, the NDHS and NNHS, which uses the Standardized Monitoring and Assessment of Relief and Transitions (SMART) methodology, calculate child anthropometric indicators for all children under five years old, whereas the ORIE evaluation reports child anthropometric indicators for all children under three years old.

2. The season in which cross-referenced surveys were conducted.

- As a number of key indicators, such as household food security, or even underweight and wasting, vary by season, so estimates from different sources may vary.
- The ORIE endline survey was conducted in July and August 2016, the NNHS 2014 was conducted from February to May, while the NNHS 2015 was conducted between July and

September 2015. The NDHS 2013 was conducted between April and May 2013, and the MICS 2011 was conducted between February and March 2011.

3. The year in which the cross-reference survey was conducted.

- Comparisons of the ORIE endline with the NDHS 2013 must be made carefully as contextual factors have very likely changed within this time gap.

4. The level of disaggregation

- Some surveys are specifically designed to provide disaggregated results for particular levels. For example, NDHS 2008 disaggregates results by state.
- Overall, the results presented in the ORIE endline report are designed to provide an endline description across the evaluation areas only. This includes 12 treatment LGAs and 12 control LGAs that span the states of Jigawa, Katsina, Kebbi and Zamfara. Hence, the present estimates are not designed to be separately representative of each of the four mentioned states. Any extrapolation of these results to each of these states, or indeed the rest of northern Nigeria, must be made with care. With that said, the comparisons presented in this section are intended to put the ORIE results into the wider context of evidence that already exists for the region.

Despite these differences, we find that the findings of this evaluation are broadly in line with the findings of other studies measuring similar indicators. This helps validate the results of this evaluation and confirms that our findings are not improbable for the current nutrition context in northern Nigeria.

Table 41 Sample size and survey characteristics

Survey	Children	Women	Households	Urban/ rural	Survey period	Year	Representativeness level
MICS 2011	0–5 years	15–49 years		Both	February–March	**2011**	National/state/regional
	n=25,192	n=30,772	n=29,077				
NDHS 2013	0–5 years	15–49 years		Both	February–May	**2013**	National/state/regional
	n=26,189	n=38,948	n=38,522				
NNHS 2014	0–5 years	15–49 years			February–May	**2014**	National/state/regional
	n=20,939	n=5,727	n=4,452				
NNHS 2015	0–5 years	15–49 years		Both	July–September	**2015**	National/state/regional
	n=20,060	n=23,688	n=25,210				
ORIE endline survey	0–3 years	15–49 years		Both	July–August	**2016**	Evaluation areas
	n=5,567	n=4,765	n=3,229				

IYCF practices

Key IYCF practices are commonly measured in a number of population-based surveys in Nigeria. The findings from the ORIE endline survey are compared to the findings of other surveys in Table 42. In summary:

- The ORIE endline survey finds that the prevalence of early initiation into breastfeeding during the first hour in the LGAs studied is 30%. Estimates from other studies in north-west Nigeria from 2011 to 2015 find a prevalence ranging from 15% to 16%. This means that, on average, ORIE finds higher estimates in its areas of study than in the broader north-west Nigeria.

- Early initiation into breastfeeding in the first 24 hours is higher than in the first hour in all studies, including ORIE. ORIE finds that, on average, 78% of children 0 to 23 months are put to breast within the first 24 hours in its areas of study. Other surveys find this rate to range from 63% to 80%. Therefore, the ORIE findings are in line with other findings in north-west Nigeria. Note that the findings in ORIE LGAs are particularly close to the ones of the NNHS 2014, and are among the highest.
- ORIE finds that, in the LGAs studied, 13% of children 0 to 5 months are exclusively breastfed. Other surveys in north-west Nigeria in 2011 and 2014 find this rate to be between 6% and 10%. The ORIE findings are therefore within a close range of the other studies presented here.
- 21% of children 6 to 23 months have the minimum dietary diversity in the ORIE areas of study. This is comparable to estimates from NNHS 2014 in the states also included in the ORIE study.

Table 42 Comparison of breastfeeding and complementary feeding with other studies

Study	Breastfeeding			Complementary feeding
	Early initiation (< 1 hour) to breastfeeding (children 0–23 months) 1/	Early initiation (< 24 hours) to breastfeeding (children 0–23 months)	Exclusive breastfeeding (children 0–5 months) 3/	Minimum dietary diversity (children 6–23 months) 4/
ORIE 2016				
Overall ORIE areas	30.3	77.8	13.1	20.8
Jigawa	19.3	75.1	9.0	21.7
Katsina	47.3	87.1	15.2	23.2
Kebbi	23.0	77.7	11.0	12.5
Zamfara	29.3	72.3	15.9	23.0
MICS 2011				
North-west Nigeria	19.9	63.1	6.2	-
Jigawa	16.6	50.9	6.4	-
Katsina	16.6	33.6	6.3	-
Kebbi	20.1	55.7	6.6	-
Zamfara	5.9	46.0	4.5	-
NDHS 2013				
North-west Nigeria	25.7	69.7	-	-
Jigawa	13.7	57.5	-	-
Katsina	18.2	82.6	-	-
Kebbi	8.3	77.4	-	-
Zamfara	12.3	47.1	-	-
NNHS 2014				
North-west Nigeria	15.4	80.4	10.3	52.6
Jigawa	0.3	65.0	-	21.3
Katsina	53.3	78.0	-	22.0
Kebbi	5.9	85.3	-	33.7
Zamfara	8.9	84.3	-	22.7
NNHS 2015				
North-west Nigeria	-	-	-	-
Jigawa	-	-	-	-
Katsina	-	-	-	-
Kebbi	-	-	-	-
Zamfara	-	-	-	-

1/ Children born in the last 0–23 months who were breastfed and put to breast within one hour of birth.
2/ Children born in the last 0–23 months who were breastfed and put to breast within 24 hours of birth.
3/ Child was fed breast milk during previous day and nothing else.
4/ Children 6–23 months who receive food from four or more food groups according to UNICEF IYCF guidelines.

In Table 43, further comparisons are made with other indicators related to IYCF practices. In summary:

- All surveys presented here find breastfeeding to be highly prevalent among their areas of study. All estimates are above the 90% threshold.
- Similarly, all surveys find very high rates of continued breastfeeding at one year within their areas of study. Estimates range from 90% in the MICS 2011 in north-west Nigeria to 98% in the ORIE quantitative evaluation in the LGAs studied.
- The ORIE report finds that continued breastfeeding at two years is about 32% in the LGAs studied. This estimate lies within the estimates of the two other surveys which measure this indicator in north-west Nigeria: namely the NNHS 2014 (26%) and the MICS 2011 (67%).
- The ORIE impact evaluation finds that 32% of children aged 6 to 23 months have the minimum meal frequency in the LGAs studied. Other surveys from 2011 to 2014 find that in north-west Nigeria between 23% and 64% of children 6 to 23 months have the minimum meal frequency. Therefore, the ORIE findings in its areas of study lies within the range of other studies.

Table 43 Comparison of breastfeeding and complementary feeding with other studies

Study	Breastfeeding			Complementary feeding			
	Child ever breastfed (children 0–23 months)	Continued breastfeeding at one year (children 12–15 months) 1/	Continued breastfeeding at two years (children 20–23 months) 2/	Introduction to solid, semi-solid and soft foods (children 6–8 months)	Received at least two milk feedings during previous day (children 6–23 months who are not currently breastfed)	Minimum meal frequency (children 6–23 months) 3/	Minimum acceptable diet (children 6–23 months) 4/
ORIE 2016							
Overall ORIE areas	99.9	97.9	31.7	65.7	4.7	31.6	9.0
Jigawa	99.8	96.3	37.9	69.1	6.9	35.5	12.3
Katsina	99.9	97.0	21.2	81.1	4.6	26.6	8.6
Kebbi	99.8	98.4	45.6	57.4	3.6	26.7	4.4
Zamfara	100.0	99.3	27.3	53.7	3.5	25.6	9.6
MICS 2011 /5							
North-west Nigeria	90.2	89.6	66.7	30.1	17.5	22.6	-
Jigawa	96.6	-	-	-	18.5	21.3	-
Katsina	61.8	-	-	-	25.0	19.8	-
Kebbi	95.7	-	-	-	16.7	22.3	-
Zamfara	89.2	-	-	-	12.2	24.9	-
NDHS 2013							
North-west Nigeria	98.3	-	-	-	10.6	64.4	-
Jigawa	98.0	-	-	-	23.8	55.3	-
Katsina	99.3	-	-	-	6.4	69.8	-
Kebbi 6/	97.9	-	-	-	(5.1)	60.1	-
Zamfara 6/	98.0	-	-	-	(5.4)	50.6	-
NNHS 2014							
North-west Nigeria	95.4	93.5	26.1	-	-	23.3	12.6
Jigawa	97.0	-	-	-	-	56.0	13.1
Katsina	88.8	-	-	-	-	39.0	9.1

Kebbi	98.6	-	-	-	-	52.2	22.6
Zamfara	97.5	-	-	-	-	58.7	13.0

Notes:

1/ Children 12–15 months of age who are currently fed breast milk and had breast milk the previous day.

2/ Children 20–23 months of age who are currently fed breast milk and had breast milk the previous day.

3/ Children 6–8 months who are breastfed and received food (excluding milk feeds) two times at least and children 9–23 months who are breastfed and received food (excluding milk feed) three times at least and children 6–23 months who are not breastfed and received food four times (including milk feeds) at least.
4/ Breastfed children 6–23 months who had at least the minimum dietary diversity and the minimum meal frequency during previous day and/or non-breastfed children 6–23 months of age who received at least two milk feedings and had at least the minimum dietary diversity (excluding milk feeds) and the minimum meal frequency during the previous day.
5/ The minimum meal frequency figures are for children aged 6–23 months who are currently not breastfed,
6/ Numbers in parenthesis are based on unweighted cases.

Vaccination status

Vaccination status is commonly measured in a number of population-based surveys in Nigeria. The findings from the ORIE endline survey are compared to the findings of other surveys in Table 44. In summary:

- The proportion of children aged 12 to 23 months having all basic vaccinations is 4% in the ORIE study (2016) in its areas of study. The MICS 2011 and NDHS 2013 find slightly higher estimates, at 8% and 10%, respectively, in north-west Nigeria.
- All surveys report measles vaccination within the 20% range in their respective areas of study. Only the NNHS 2014 finds a higher rate of 42%.
- Coverage of polio vaccination at birth in the ORIE study was found to be 34%. Other surveys covering a period from 2011 to 2015 find estimates ranging from 20% to 26%. This means that polio vaccination at birth was slightly higher in the ORIE areas of study in 2016 than in the other studies reporting this indicator for north-west Nigeria.
- More than 80% of children aged 12 to 23 months were found to have received polio vaccination (polio 1, 2 or 3) in the ORIE areas of study in 2016. The MICS 2011 and NDHS 2013 find slightly lower estimates in north-west Nigeria: polio vaccination (1, 2 or 3) ranges from 60% to 30% in the MICS report, while it ranges from 77% to 61% in the NDHS 2013.
- The ORIE impact evaluation finds that about 25% of children aged 12–23 months are vaccinated against BCG on average in its LGAs of interest. This is within the range of estimates found by the MICS 2011 (32%) and NDHS 2013 (22%) in north-west Nigeria.
- The figures reporting the percentage of DTP/PENTA vaccinations (ranging from 1 to 3) are lower in ORIE 2016 compared to the other surveys presented here. ORIE finds that between 5% and 15% of children aged 12–23 months have received at least one shot of DPT/PENTA vaccines in the LGAs studied, while these estimates range from 13% to 36% in the other surveys for the whole of north-west Nigeria.

Table 44 Comparison of vaccination status of children (12–23 months) with other studies

	Vaccination										
	Proportion of children not having any vaccination (12–23 months)	Proportion of children aged 12–23 months having all basic vaccination /1	Measles (12–23 months)	Polio 0	Polio 1	Polio 2	Polio 3	BCG	DTP1 / Penta1 (12–23 months)	DTP2 / Penta2 (12–23 months)	DTP3 / Penta3 (12–23 months)
ORIE 2016											
North-west Nigeria	-	4.2	29.8	34.4	84.8	82.7	80.3	24.6	15.2	6.4	5.3
Jigawa	-	3.3	33.3	36.2	79.9	77.8	75.7	26.6	17.7	6.2	5.1

Katsina	-	10.4	44.4	58.6	87.7	85	82.8	40.4	25.6	13.1	10.4
Kebbi	-	1.1	26.0	26.0	93.6	91.6	87.6	12.4	6.8	2.5	1.8
Zamfara	-	1.9	17.6	18.3	81.6	79.9	77.8	21.3	10.2	3.6	3.4
MICS 2011											
North-west Nigeria	36.0	7.5	28.8	19.3	61.2	50.1	31.3	32.1	27.8	21.5	13.2
Jigawa	51.2	4.6	22.3	13.2	41.9	34.2	21.6	26.9	20.7	12.4	7.2
Katsina	30.1	9.2	38	26.8	68.6	58.3	37.8	31.5	27.6	19.7	11.4
Kebbi	27.6	4.3	19.5	15.5	68.6	58.3	37.8	20.4	11.7	10.6	6.5
Zamfara	38.1	1.8	11.4	7.7	60.6	54.1	40.6	21.4	15.1	4.5	3.7
NDHS 2013											
North-west Nigeria	20.8	9.6	22.3	26.4	77.2	71.8	61.1	21.7	22.2	18.1	13.9
Jigawa	29.9	3.6	10.9	23.2	67.2	59.3	49.1	20.4	20.9	12.7	7.0
Katsina	4.9	8.7	42.9	25.7	94.1	88.7	81.3	21.9	23.3	17.9	14.6
Kebbi	5.8	2.8	3.2	38.5	92.8	90.0	80.5	5.2	4.9	4.2	2.8
Zamfara	28.6	2.1	7.9	7.6	69.1	63.7	53.9	10.5	12.2	8.1	5.6
NNHS 2014											
North-west Nigeria	-	-	42.5						35.8	27.2	17.9
Jigawa	-	-	38.2						30.6	19.4	6.5
Katsina	-	-	44.4						35.9	30.3	20.4
Kebbi	-	-	48.9						15.4	11.0	7.1
Zamfara	-	-	18.6						12.4	8.1	5.0
NNHS 2015											
North-west Nigeria	-	-	21.5						30.9	23.5	16.9
Jigawa	-	-	29.7						37.6	27.2	17.3
Katsina	-	-	21.0						35.3	28.7	19.8
Kebbi	-	-	14.7						21.1	15.2	7.8
Zamfara	-	-	7.0						11.4	9.2	5.9

Notes:
1/ Basic vaccination includes polio (four shots), DPT (three shots), BCG, measles, yellow fever, HepB (four shots).
NDHS and ORIE do NOT include yellow fever and HepB.

Child anthropometric and malnutrition analysis

Table 45 shows that recent studies on nutrition in Nigeria provided similar but somewhat diverging and varying measures of malnutrition among children in northern Nigeria. In summary:

- Based on our survey data, ORIE estimates a prevalence of stunting of about 51% across the LGAs surveyed in this study. For north-west Nigeria, and between 2011 and 2015, estimates of stunting vary between 50% and 56%. This implies two things: first, in terms of stunting, the estimates derived from areas surveyed in the present study plausibly lie within estimates presented in other recent surveys implemented in the region. Note that the fact that the ORIE estimate is at the lower end of this range might be due to the fact that we are surveying younger children, among whom stunting is generally less prevalent. Second, these findings imply that stunting continues to be a severe public health problem in northern Nigeria.

- Across the areas surveyed in this evaluation, we estimate a prevalence of underweight of 38%. Estimates from other surveys in the area (although again with a different base population) vary between 31% and 38%.

Table 45 Comparison of anthropometric estimates with other studies

Study	Base population	Prevalence			
		Stunting		Underweight	
		Stunting (HAZ<-2SD)	Moderate stunting (HAZ<-2 & >=3SD)	Underweight (WAZ<-2SD)	Moderate underweight (WAZ<-2SD)
ORIE 2016					
North-west Nigeria	Children aged 0–35 months	50.9	-	38.2	-
Jigawa		54.7	-	42.1	-
Katsina		50.2	-	39.9	-
Kebbi		49.4	-	36.5	-
Zamfara		49.6	-	34.6	-
MICS 2011					
North-west Nigeria	Children aged 0–59 months	53.8	-	38.4	-
Jigawa		58.8	-	43.8	-
Katsina		61.9	-	44.8	-
Kebbi		53.9	-	43.4	-
Zamfara		61.7	-	47.5	-
NDHS 2013					
North-west Nigeria	Children aged 0–59 months	54.8	-	27.1	-
Jigawa		59.0	-	17.0	-
Katsina		58.5	-	24.3	-
Kebbi		60.6	-	18.1	-
Zamfara		55.9	-	16.2	-
NNHS 2014					
North-west Nigeria	Children aged 0–59 months	50.4	29.6	32.7	23.5
Jigawa		55.5	32.3	40.2	26.7
Katsina		58.1	30.2	33.2	24.6
Kebbi		46.4	30.2	28.1	19.2
Zamfara		49.9	27.4	30.6	23.0
NNHS 2015					
North-west Nigeria	Children aged 0–59 months	55.9	31.4	31.6	22.3
Jigawa		63.4	31.4	40.6	27.3
Katsina		58.2	31.9	33.9	23.5
Kebbi		58.3	33.2	33.5	23.9
Zamfara		57.5	32.5	28.0	21.4

Table 46 compares estimates of acute malnutrition (wasting) between ORIE and other surveys. In summary:

- Based on our survey data, ORIE estimates a prevalence of global acute malnutrition (GAM) ranging from 14% to 17% in the LGAs surveyed in this study, depending on the measure used. For north-west Nigeria, and between 2011 and 2015, estimates of GAM vary between 7% and 27%. This implies that in terms of GAM the estimates derived from areas surveyed in this present study plausibly lie within estimates presented in other recent surveys implemented in the region. In a similar way, ORIE estimates a prevalence of moderate acute malnutrition

(MAM) ranging from 9% to 11%, depending on the measure used. Other surveys presented here find estimates ranging from 5 to 8%. While the estimates in the ORIE LGAs are slightly higher, the findings seem reasonably comparable to findings in north-west Nigeria.

- The ORIE impact evaluation finds SAM prevalence to range from 4% to 7% in the LGAs studied and depending on the measure used. Other surveys spanning 2011 to 2015 and covering north-west Nigeria find estimates ranging from 2% to 15%. There are two things to note: first, the ORIE estimates lie within the estimates of other surveys measuring SAM. Second, ORIE is the only study that allows estimation of SAM through different measures. While ORIE estimates using only z-scores are lower than surveys using this measure (MICS 2011), ORIE estimates are higher, when using a combination of MUAC, oedema and z-scores, than the other surveys (NNHS 2014 and 2015). Finally, this also implies that SAM remains an important health issue in north-west Nigeria in general.

Table 46 Comparison of acute malnutrition with other studies

Study	Base population	GAM			MAM		SAM			
		(WHZ <-2)	(WHZ <-2 and/or oedema)	(MUAC <125 and/or oedema)	(WHZ <-2 >= -3 no oedema)	(MUAC<125 >=115, no oedema)	(WHZ<-3 and/or oedema)	(MUAC <115 and/or oedema)	SAM (MUAC<115)	WHZ<-3SD
ORIE 2016										
Overall ORIE areas	Children aged 6–35 months	17.3	18.2	14.2	11.2	8.7	7.1	5.5	4.4	5.8
Jigawa		20.0	21.4	16.1	11.8	10.3	9.6	5.9	3.6	7.5
Katsina		18.7	19.5	13.5	12.5	8.5	7.0	5.0	4.1	5.9
Kebbi		16.6	17	13.7	11.8	9.2	5.2	4.5	3.8	4.6
Zamfara		14.4	15.3	13.7	9.1	7.3	6.2	6.4	5.7	5.0
MICS 2011										
North-west Nigeria	Children aged 0–59 months	13.9	-	-	-	-	-	-	-	4.9
Jigawa		14.3	-	-	-	-	-	-	-	6.6
Katsina		14.7	-	-	-	-	-	-	-	5.1
Kebbi		18.2	-	-	-	-	-	-	-	5.7
Zamfara		17.5	-	-	-	-	-	-	-	6.7
NDHS 2013										
North-west Nigeria	Children aged 0–59 months	27.1	-	-	-	-	-	-	-	15.3
Jigawa		17.0	-	-	-	-	-	-	-	7.8
Katsina		24.3	-	-	-	-	-	-	-	12
Kebbi		18.1	-	-	-	-	-	-	-	9.4
Zamfara		16.2	-	-	-	-	-	-	-	6.1
NNHS 2014										

North-west Nigeria	Children aged 6–59 months	-	10.5	6.2	8.0	4.6	2.5	1.6	-	-
Jigawa		-	17.8	7.6	14.1	6.3	3.7	1.3	-	-
Katsina		-	7.3	9.1	5.8	6.2	1.6	2.9	-	-
Kebbi		-	9.9	7.6	7.9	6.2	2.0	1.4	-	-
Zamfara		-	8.7	6.1	7.3	4.8	1.4	1.3	-	-
NNHS 2015										
North-west Nigeria	Children aged 6–59 months	-	10.2	10.3	7.1	7.9	3.1	2.4	-	-
Jigawa		-	11.9	12.5	10.2	10.3	1.7	2.2	-	-
Katsina		-	10.3	8.1	8.2	5.6	2.1	2.5	-	-
Kebbi		-	9.3	12.8	6.4	10	2.9	2.8	-	-
Zamfara		-	7.1	11.2	6.2	8.9	0.9	2.3	-	-

Uptake of Vitamin A supplements, ORS and deworming medication

The results presented in Table 47 show the differing situations in north-west Nigeria regarding the uptake of Vitamin A, ORS and deworming medication. In summary:

- The ORIE study finds that 23% of children aged 6–35 months have received Vitamin A supplements in the last six months in the LGAs studied. Other surveys in north-west Nigeria, and between 2011 and 2015, find estimates ranging from 16% to 55%. This implies that the ORIE findings are well within the range of estimates from surveys collecting this information in north-west Nigeria.
- Similarly, the ORIE impact evaluation finds 38% of children aged 0–35 months have had diarrhoea in the two weeks preceding the survey and that 24% received ORS treatment for it. Other surveys presented here and collecting this information find that between 9% and 29% of children aged 0–59 months have had diarrhoea in the last two weeks preceding the survey and that between 14% and 38% of them received ORS treatment for it. Again, while it is important to note the difference in age range and in geographical span of these indicators between ORIE and other studies, it seems that the ORIE findings lie comfortably within the estimates of the other surveys presented here.
- Finally, based on the ORIE data, 11% of children aged 12 to 35 months in the ORIE LGAs studied have received deworming medication in the last six months. The NNHS surveys (2014 and 2015) find that 9% of children aged 12 to 59 months have received deworming medication in north-west Nigeria. While the ORIE findings are slightly lower, they remain within a close range of these other survey estimates.

Table 47 Comparison of Vitamin A, ORS and deworming medication intake by children with other studies

Survey	Vitamin A	Diarrhoea		Deworming	
	Vitamin A supplement in the last 6 months (children 6–59 months) /1	Diarrhoea in the last 2 weeks (children 0–59 months) /2	ORS treatment for diarrhoea (children with diarrhoea in the last 2 weeks)	Deworming medication in the last 6 months (children 6–59 months) /1	Deworming medication in the last 6 months (children 12–59 months) /4
ORIE 2016					
Overall ORIE areas	22.9	37.8	24.0	10.6	11.4
Jigawa	15.7	38.2	24.5	9.6	10.1
Katsina	37.5	34.6	24.9	19.9	21.7
Kebbi	23.2	34.1	22.7	9.9	10.4
Zamfara	16.3	42.3	23.7	4.0	4.8
MICS 2011					
North-west Nigeria	55.4	20.1	26.0	-	-
Jigawa	64.6	-	-	-	-
Katsina	71.9	-	-	-	-
Kebbi	68.4	-	-	-	-
Zamfara	41.8	-	-	-	-
NDHS 2013					
North-west Nigeria	26.1	9.2	33.7	9.0	-
Jigawa	17.7	14.8	-	6.7	-
Katsina	77.1	7.7	-	45.2	-
Kebbi	10.7	13.6	-	1.4	-

Zamfara	24.5	6.0	-	0.5	-
NNHS 2014					
North-west Nigeria	36.3	29.0	13.8	-	8.6
Jigawa	32.3	33.6	18.5	-	18.7
Katsina	56.9	26.6	13.5	-	15.0
Kebbi	15.2	33.1	0.4	-	0.8
Zamfara	46.7	33.0	3.2	-	1.4
NNHS 2015					
North-west Nigeria	15.5	24.3	22.4	-	9.2
Jigawa	16.3	24.1	38.8	-	16.5
Katsina	19.0	22.2	15.1	-	8.6
Kebbi	20.2	36.7	9.7	-	12.2
Zamfara	1.4	24.7	8.6	-	0.8

Notes:

1/ the age group in the ORIE surveywas children aged 6–35 months.

2/ the age group in the ORIE surveywas children aged 0–35 months.

3/ the age group in the ORIE surveywas children aged 12–35 months. The age group is different to that in Volume 1 in order to facilitate better comparison with other surveys.

References

Attané, I. and Guilmoto, C.Z., eds. (2007) 'Fertility Watering the Neighbour's Garden. The Growing Female Deficit in Asia', CICRED, Paris, p. 416.

Belloni, A., Chernozhukov, V. and Hansen, C. (2014) 'High-Dimensional Methods and Inference on Structural and Treatment Effects', *Journal of Economic Perspectives* 28(2): 29–50.

Black, R.E., Allen, L.H., Bhutta, Z.A., Caulfield, L.E., de Onis, M., Ezzati. M., Mathers, C., Rivera, J., for the Maternal and Child Undernutrition Study Group 2008 (2008) 'Maternal and child undernutrition: global and regional exposures and health consequences', *The Lancet* 371(9608), pp. 243–60. Cited in DFID (2009) 'The neglected crisis of undernutrition: Evidence for action', p 16.

Caliendo, M. and Kopeinig, S. (2005) 'Some practical guidance for the implantation of propensity score matching', IZA Discussion Papers, No. 1588.

Cogill, B. (2003) 'Anthropometric Indicators Measurement Guide', Washington DC: Food and

Colin Cameron, A., Gelbach, J.B., Miller, D.L. (2008) 'Bootstrap-based improvements for inference with clustered errors', *The Review of Economics and Statistics*, 90(3): 414–427.

Deitchler, M., T. Ballard, A. Swindale, and J. Coates (2011) 'Introducing a Simple Measure of Household Hunger for Cross-Cultural Use', Washington DC: Food and Nutrition Technical Assistance II Project.

DFID (2011) 'Improving newborn and child nutrition in Northern Nigeria: Business Case', London.

Eldridge, S.M., D. Ashby, and S. Kerry (2006) 'Sample size for cluster randomized trials: effect of coefficient of variation of cluster size and analysis method', *International Journal of Epidemiology* 35, pp. 1292–1300.

Gabler, S., S. Hader, P. Lynn (2006) 'Design effects for multiple design samples', *Survey Methodology* 32(1), pp. 115–120.

Grosh, M. E. and J. Munoz (1996) 'A manual for planning and implementing the living standards measurement study survey' in Living Standards Measurement Study Working Papers Working Paper No. 126, Washington: The World Bank.

Hemming, K., A.J. Girling, A.J. Slitch, J. Marsh and R.J. Lilford (2011) 'Sample size calculations for cluster randomised controlled trials with a fixed number of clusters', *BMC Medical Research Methodology* 11(102), pp. 1–11.

Imbens, G. and Rubin, D. (2015) *Causal Inference for Statistics, Social, and Biomedical Sciences*. Cambridge University Press.

Kostermans, C.P. (1994) *Assessing the quality of anthropometric data: background and illustrated guidelines for survey managers* Living standards measurement study (LSMS) working paper; no. LSM 101. Washington, DC : The World Bank. http://documents.worldbank.org/curated/en/251771468765256658/Assessing-the-quality-of-anthropometric-data-background-and-illustrated-guidelines-for-survey-managers

Limpert, E., Stahel, W.A. and Abbt, M. (2001) 'Log-Normal Distributions Across the Sciences: Keys and Clues' *Bioscience* 51:341–52.

Mei, Z. and Grummer-Strawn, L. (2007) 'Standard Deviation of Anthropometric Z-scores as a Data Quality Assessment Tool Using the 2006 WHO Growth Standards: A Cross Country Analysis', *Bulletin of the World Health Organization* 85 (6), pp. 441–448. www.scielosp.org/pdf/bwho/v85n6/a10v85n6.pdf.

National Planning Commission (2004) 'National Plan of Action on Food and Nutrition in Nigeria'. Available at: https://extranet.who.int/nutrition/gina/sites/default/files/NGA%202005%20National%20Plan%20of%20Action%20on%20Food%20and%20Nutrition.pdf.

National Population Commission (NPC) Nigeria and ICF International (2013) 'Nigeria Demographic and Health Survey 2013: Preliminary Report', Abuja, Nigeria: National Population Commission and ICF Macro. [Cited as NDHS 2013].

NDHS (2008) 'Nigeria Demographic and Health Survey 2008', Abuja: NPC. Available at: www.dhsprogram.com/pubs/pdf/FR222/FR222.pdf.

NNHS (2014) 'National Nutrition and Health Survey (NNHS) – report on the nutrition and health situation of Nigeria, 2014'.

NNHS (2015) 'National Nutrition and Health Survey (NNHS) – report on the nutrition and health situation of Nigeria, 2015'.

Nutrition Technical Assistance (FANTA) Project, FHI 360.

Omilola, B. (2010) 'Patterns and trends of child and maternal nutrition inequalities in Nigeria', IFPRI Discussion Paper 968. Washington, DC: International Food Policy Research Institute (IFPRI). Available at: http://ebrary.ifpri.org/cdm/ref/collection/p15738coll2/id/1578.

Rosenbaum, P.R. and Rubin, D.B. (1985) 'Constructing a control group using multivariate matched sampling methods that incorporate the propensity score', *The American Statistician* 39:33–38.

Rubin, D. (2001) 'Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation', *Health Services & Outcomes Research Methodology* 2, 169–188.

Shiyuan C., M. Schreiner and G. Woller (2008) 'A Simple Poverty Scorecard for Nigeria', unpublished, www.microfinance.com/English/Papers/Scoring_Poverty_Nigeria_EN_2003.pdf, [Accessed 11 May 2017].

Silventoinen, Karri (2003) 'Determinants of variation in adult body height', *Journal of Biosocial Science* 35 (2):263–285.

Snedecor, G.W. and Cochran, W.G. (1989) *Statistical Methods*. Ames, IA: Iowa University.

Standardized Monitoring and Assessment of Relief and Transitions (2012) 'Sampling Methods and Sample Size Calculation for the SMART Methodology', June 2012, http://smartmethodology.org/documents/sampling/Sampling_Paper_June_2012.pdf, [Accessed 12 May 2017].

Tanner, J.M. (1981) *A History of the Study of Human Growth*. Cambridge, England: Cambridge University Press.

UNHCR (2013). 'Standardised Expanded Nutrition Surveys (SENS) Guidelines for Refugee Populations, Module 1: Anthropometry and Health'.

Vincenty, T. (1975) 'Direct and inverse solutions of geodesics on the ellipsoid with application of nested equations', *Survey Review* 22(176): 88–93.

Visram, A., Jasper, P., Moore, L., Adegoke, F., Kveder, A., Arif, S. and Ward, P. (2014) 'ORIE Nigeria: Quantitative Impact Evaluation. Baseline Report', Oxford: ORIE, www.cmamforum.org/Pool/Resources/ORIE-quantitative-full-baseline-report-Nigeria2014.pdf.

WHO (1995) 'Physical Status: The Use and Interpretation of Anthropometry', in WHO Technical Report Series (854), pp. 1–440. http://whqlibdoc.who.int/trs/WHO_TRS_854.pdf. [Accessed 15 August 2014].

WHO (1995) 'Physical Status: The Use and Interpretation of Anthropometry' in WHO Technical Report Series (854), pp. 1–440. http://whqlibdoc.who.int/trs/WHO_TRS_854.pdf.

WHO (2003) 'The Management of Nutrition in Major Emergencies', Geneva, WHO.

WHO (2008) 'Indicators for Assessing Infant and Young Child Feeding Practices: Part 1 Definitions', Geneva, Switzerland: WHO.

WHO/UNICEF Joint Monitoring Programme for Water Supply and Sanitation (2012) 'Progress on Drinking Water and Sanitation: 2012 Update'. UNICEF and World Health Organization, USA

www.wssinfo.org/fileadmin/user_upload/resources/JMP-report-2012-en.pdf. [Accessed 15/03/2017].

Wooldridge, J. (2012). *Introductory Econometrics*, South-Western Cengage Learning, 5th Edition, Mason: USA.

Annex A Terms of Reference

Terms of Reference for Nutrition programme: Operational research and impact evaluation (ORIE)

Introduction
DFID is seeking proposals from research and academic institutions to plan, manage and implement an operational research and impact evaluation (ORIE) component of a large programme to tackle undernutrition in Northern Nigeria. The £ 50 million overarching Improving Maternal, New-born and Child Nutrition in Northern Nigeria programme will be implemented by Government of Nigeria, UNICEF, Save the Children and Action Against Hunger across 5 states, over 6 years and focuses on scaling up direct nutrition interventions known to be effective in tackling undernutrition and reducing mortality. The ORIE component would work closely with those implementing the programme, and also link with DFID funded research programmes in Nutrition. In addition, the selected supplier will be responsible for tracking and coordinating for onward transmission to DFID, progress reports from all the implementing partners delivering programme outputs.

Background:
There is a high prevalence of undernutrition amongst children under five in Nigeria, and particularly in the north: in this part of the country, half of all children under five are stunted, and one in five suffers from acute malnutrition. This has profound implications for health and for human development, and presents a major obstacle to attainment of the Millennium Development Goals in the country as well as globally.

The nutrition programme will provide a number of evidence-based, highly cost-effective direct interventions for the prevention and treatment of malnutrition, including community-based management of acute malnutrition (CMAM), vitamin A supplementation and deworming, and promotion of improved infant and young child feeding (IYCF) practices. Delivery of these interventions will be used to raise the profile of undernutrition on the political and development agenda in Nigeria, and to enhance the commitment and capacity of government and others to address its immediate, basic and underlying causes. The log frame for the nutrition programme can be found in annex 1.

Objective
The objective of the ORIE component is to determine the impact of DFID Nigeria's Nutrition programme and address key evidence gaps on solutions to under nutrition in Northern Nigeria.

Methodology and Scope of Work
The ORIE component should focus on four areas (outlined below). Proposals should focus on: further developing these ideas and how to take them forward; as well as on how to plan and implement the component in support of the wider programme.
Major outputs of the ORIE component:

1. **Evidence on the best approaches to scaling up interventions known to be effective in reducing undernutrition in the Northern Nigeria context.** This will include designing and implementing studies to test different implementation strategies, and ensuring the findings are used to adjust programme design in order to maximise impact and programme efficiency. This may include for example: testing various

delivery platforms for therapeutic zinc supplementation; testing approaches to developing and maintaining a strong community based work force to support the delivery of direct nutrition services; testing approaches for empowering women to demand better health services from ward and local government; and developing methods for measuring value for money of various implementation strategies.

2. **Evaluation of the cost effectiveness and DFID-attributable impact of the nutrition programme.** This will include designing an evaluation strategy which allows DFID to be fully accountable for funds committed to the programme and which generates a clear evidence base on what the programme has achieved in accordance with the log frame (Annex 1). The design should include consideration of elements of the programme which already have a strong evidence of impact and those which do not and propose a strategy which delivers best value for money. The programme implementers (UNICEF, Save the Children and Action against Hunger) will design the programme monitoring component to support the impact evaluation plan. The evaluation should also allow the cost effectiveness of the programme to be judged.

3. **Evidence on complementary solutions to stunting.** Undertaking a cohort study if the literature review during inception indicates that it will provide valuable new information.

4. **Dissemination and uptake of evidence.** This will include publication of results from outputs 1-3 in a range of products suitable for programme partners and policy makers in Nigeria and beyond. It will also include publications for peer-reviewed journals. A strategy for dissemination will include meetings, events and conferences in Nigeria and (where appropriate) beyond.

Recipient
The recipient of this work will be the Government of Nigeria, the Government of Jigawa, Katsina, Kebbi, Yoba and Zamfara States and DFID Nigeria Abuja and Kano Offices

Timeframe
The timing of the Operational Research and Impact Evaluation work will start as soon as possible after finalising the contract details but the aim is to commence the service no later than the end of February 2012. DFID intends to let the contract for an initial period of 5 years, with a possible extension of up to 1 year.

Reporting
Reporting will be direct to DFID Nigeria with a copy to the Project Management Board.

DFID co-ordination
DFID Nigeria is the sole funder of the ORIE component and the Health Adviser for Northern Nigeria will be responsible for ensuring the component is implemented according to plan.

Further Background

What need are we trying to address?
One million children under five die every year in Nigeria, 35% of them due to causes attributed to malnutrition. This makes Nigeria one of the six countries that accounts for half of all child deaths from malnutrition worldwide. In the north, half of all children under five are stunted, and one in five

suffers from acute malnutrition. This has profound implications for health and for human development, and presents a major obstacle to attainment of Millennium Development Goals in the country and globally. To date, the Nigerian government has not provided the necessary leadership or response to the crisis. Coupled with this, is a weak and fragmented health system which is unable to provide the most basic, cost-effective services for the prevention and management of common health problems. Primary health care level remains the weakest link in effective health delivery.

What will we do to tackle this problem?

The programme will deliver a number of evidence-based, highly cost-effective direct interventions for the prevention and treatment of malnutrition, including community-based management of acute malnutrition (CMAM), vitamin A supplementation and deworming, and promotion of improved infant and young child feeding (IYCF) practices. The scaled up delivery will be used to raise the political profile of undernutrition in Nigeria and leverage government to coordinate and fund nutrition programmes. Independent operational research will examine the wider determinants and structural barriers of undernutrition. Impact evaluation will measure progress, quality and advise on critical elements required for a sustainable strategy.

Who will be implementing the support we provide?

A UNICEF and an INGO consortium of Save the Children (SC UK) and Action against Hunger / Action against Hunger will deliver the interventions. Operational research and impact evaluation will be conducted by independent nutrition researchers and evaluation experts.

Annex B Inception Report

B.1 Volume 1 (excerpts)

B.1.1 Impact evaluation (IE)

B.1.1.1 Overview of the impact evaluation

The primary rationale of the impact evaluation work stream is to provide an independent assessment of the overall impact of the WINNN programme. It is important to clarify that the impact evaluation will not set out to determine the impact of individual WINNN Programme technologies. There already exists an extensive evidence-base for the individual-level impact of a number of the different nutritional interventions that will be deployed in the WINNN programme. The design of the WINNN Programme is such that interventions are integrated within existing routine health services. The objective is to strengthen service provision directly and to improve government ownership and finance of these services to provide a basis for long term sustainability. The first two outputs are focussed on strengthening primary care services and nutritional education in three focal LGAs per state. It is taken that the objective is to provide coverage for the entire population of these LGAs, with a view to further roll-out by the state government in future. The third and fourth outputs are state-wide and (for the fourth) national

As a result, the impact assessment will:

1. Assess the impact of those WINNN interventions focussed in particular LGAs using quantitative estimates of impact on population-based, LGA-wide indicators of nutritional behaviours and nutritional status (this will include indicators at higher levels in the WINNN logframe, particularly anthropometric status); and

2. Use qualitative analysis, drawing where possible on quantitative data, to make an assessment of the impact of WINNN on state-wide indicators and on measures of policy change and government co-financing for state- and LGA-level indicators.

As such, the impact evaluation will use a quasi-experimental design to assess the collective impact of the WINNN Programme interventions for Outputs 1 and 2. The impact of WINNN Programme interventions for Outputs 3 and 4 will be assessed using mixed evaluative methods. Table 3 provides a summary of the methods that will be used to investigate each component of the WINNN Programme.

Table 1: Summary of impact evaluation research methods by WINNN Programme

WINNN Programme Component		**Research method**
Impact	Improved nutritional status of children under five in Northern Nigeria	Quantitative impact analysis
Outcome	Delivery of nutrition interventions through routine health services, funded by the Government of Nigeria	Qualitative impact assessment
Output 1	Delivery of effective treatment for severe acute malnutrition through local health	Quantitative impact analysis

	systems in selected states and local government areas in Northern Nigeria	Qualitative impact assessment
Output 2	Delivery of effective infant and young children feeding interventions in selected states and local government area in Northern Nigeria	Quantitative impact analysis Qualitative impact assessment
Output 3	Integration of micronutrient interventions into routine primary health services	Qualitative impact assessment
Output 4	Strengthening of nutrition coordination and planning mechanisms at the national and state level	Qualitative impact assessment

The sections that follow provide a brief overview of the quantitative and qualitative methods that will be used in the impact evaluation.

B.1.2 Quantitative data – household survey

B.1.2.1 Rationale and objectives of the quantitative impact evaluation

The main objective of the quantitative component of the impact evaluation work stream is to provide an independent assessment of the combined impact of the Community Management of Acute Malnutrition programme (Output 1) and the Infant and Young Child Feeding programme (Output 2). Note that since all four WINNN programmes are implemented together, the quantitative evaluation constitutes an evaluation of these two programmes in the presence of the activities for outputs 3 and 4, i.e. the integration of micro-nutrient interventions into routine primary health care services, and the strengthening of nutrition coordination and planning mechanisms.

B.1.2.2 Approach and methods

A quasi-experimental design with difference-in-difference estimates will be used to assess the collective impact of the WINNN programme interventions for Outputs 1 and 2. The evaluation model consists of two groups, a treatment group and a comparison group. These interventions will be implemented in the treatment group while there will be no intervention in the comparison group. For every group of intervention LGAs in a State, comparison-group LGAs will be purposively selected by ORIE to be as similar as possible to the intervention LGAs. In both groups, quantitative data will be collected in a random sample of households at baseline (planned for the first quarter of 2013) and again at follow-up three years later. Because the treatment areas were not selected randomly, survey results will be representative of, or generalizable to, only the WINNN programme areas, not to all of Northern Nigeria. We will use the knowledge of our partners and the WINNN IPs to contextualise the results and suggest their applicability beyond WINNN programme areas. Econometric modelling to control for differences between the treatment and control groups which may co-determine the impact indicators will be used.

While the WINNN Programme logframe will guide the impact evaluation, a specific Theory of Change Model will be developed to determine specific indicators. This Theory of Change Model will be developed in collaboration with the Operations Research work stream as it will also signal important questions for operations research. Table 2 provides a broad overview of the possible areas of impact and related indicators.

Table 2: Possible areas of impact and indicators for the quantitative impact evaluation

WINNN Programme Component	Hypothesis	Impact indicators
OUTPUT 1 (CMAM programme)	Timely detection of severe acute malnutrition (SAM) in a community and provision of treatment with RUTF for children without complications and combined with facility-based approaches for children with complications will reduce the prevalence of SAM	Under 5 prevalence of moderate and severe wasting
OUTPUT 2 (IYCF interventions)	Community-based approaches to deliver effective feeding practices including breastfeeding interventions will improve breastfeeding, weaning and feeding practices	Under 5 prevalence of wasting Under 5 prevalence of underweight Under 5 prevalence of stunting Proportion of infants 0-6 months exclusively breastfed Proportion of children 6-23 months receiving foods from four or more food groups supplementation in the last 6 months Proportion of women who took iron supplementation for at least 90 days during pregnancy for latest birth Early initiation of breastfeeding

A competitive tender process has been launched to identify the best-qualified survey organisation in Nigeria to collect data for the quantitative impact evaluation. The survey organisation will provide personnel for all aspects of data collection and entry, from field managers and enumerators to data managers and data entry clerks. Their work will be closely supervised by OPM and national partners.

B.2 Volume 2 (excerpts)

Quantitative component: household survey

Quantitative evaluation methods

Quantitative evaluation model

A quasi-experimental design is proposed to assess the collective impact of the WINNN Programme interventions for Outputs 1 and 2. The evaluation model consists of two groups, a Treatment Group and a Control group. All components of the WINNN Programme related to Output 1 and 2 will be implemented in the Treatment Group while there will be no intervention by the WINNN Programme in the Control Group. In both groups, quantitative data will be collected at baseline (before the intervention is implemented), and again at follow-up which will be two years later. Data will be collected in a sample of health facilities and households.

Selection of the Treatment Group

Due to the nature of the WINNN Programme, the Treatment Group has already been identified in collaboration with Government stakeholders and DFID. Within each State, the WINNN Programme will implement Output 1 (CMAM programme) and Output 2 (IYCF interventions) activities in three purposively selected LGAs. These LGAs were selected between WINNN and the State Governments. Within each LGA, the WINNN Programme will progressively roll-out their interventions to cover a minimum of 10 wards. It is the understanding of the ORIE quantitative impact evaluation team that ORIE will not be able to influence the selection of States, LGAs or Wards in which the WINNN Programme will be implemented.

Activities related to Output 3 (micronutrients and deworming / MNCHW events) and Output 4 (government coordination) will be implemented across the entire State. The impact of these outputs will primarily be assessed using qualitative methods and are therefore discussed in more depth in a later section.

Selection of the Control Group

As ORIE was not able to influence the selection of the Treatment Group, it will be important to pay special consideration to the selection of the Control Group to strengthen the robustness of the impact analysis. There are three options for the selection of the Control Group; the first is to look outside of the selected State that WINNN will be implementing its programmes and the second option is to look within the selected State, either at the level of the LGA or the Ward.

While the first option (looking for a control group outside of the selected WINNN State) may seem like a simple solution, one must consider the uniqueness of each State in Nigeria. The health system in each State is run independently from one another and significant differences in the strength of the healthcare system, access to healthcare and burden of disease exist between States. Therefore, the Control Group must be selected within each selected WINNN State.

Two options present themselves when selecting the Control Group from within the WINNN State, Control Groups at the level of the Ward or at the level of the LGA. Because the CMAM and IYCF interventions are non-targeted supply-side interventions, selection of the control group at the level of the Ward might risk 'contamination' as WINNN Programme beneficiaries in control Wards would easily be able to seek out treatment or exposure to the WINNN interventions in treatment Wards. Therefore, we propose selecting the control group from amongst non-WINNN LGAs.

The gold-standard in impact evaluation is to randomly allocate treatment and control status to produce two evaluation groups that have a high probability of being statistically identical – so long as the number of study locations is sufficiently large. Furthermore, random assignment will minimise any biases in the assessment of impact due to the effect of other health and social welfare interventions that are concurrently implemented in the State. In this case, it is not possible to randomly assign Treatment LGAs because these have been pre-specified by the WINNN Programme. As such, it will be important to purposively select Control Group LGAs so that they are as similar to the Treatment Group LGAs as possible. To this end, a mapping and information gathering exercise of all LGAs will be conducted. This exercise will provide important information to allow for the selection of the most suitable non-WINNN LGAs within the State. The following are dimensions that will be considered in the selection of the Control Group LGAs, in so far as the information can be obtained[7]:

[7] The availability and quality of data on LGAs is often very poor

- Socio-economic and demographic indicators
- Access to health care and burden of disease
- Proximity to WINNN-LGAs
    - Due to the nature of the WINNN interventions, it is expected that there will be spill-over effects into adjacent LGAs. Patients from LGAs adjacent to the WINNN-LGA may travel to seek out treatment. Therefore, LGAs adjacent to WINNN-LGAs will not be selected to form the Control Group
- Presence of other health and nutrition interventions
    - There are a number of health and nutrition interventions being implemented in parallel to the WINNN Programme. As such, the mapping exercise will compile a list of all current and future health and nutrition interventions in the State. With this information, every effort will be made to select a Control Group LGA that is as similar to the Treatment Group LGAs as possible, including implementation of 'third party' interventions.

Following receipt of the national consultant's report, the ORIE quantitative impact evaluation team will be able to propose a set of Control Group LGAs. Depending on the number of suitable Control Group LGAs within a given State, propensity score matching techniques may be used to make the final selection. It is important to note here that Control Group LGAs will need to be established with State governments before the study can begin to ensure they will remain un-contaminated throughout the study period.

Difference-in-differences

Despite having a purposively selected Control Group LGAs, the evaluation may still result in evaluation groups that are not statistically balanced. Therefore, difference-in-difference methods will be used to account for any predominant trends that may bias the assessment of impact.

Using a difference-in-difference approach, the impact using the evaluation model presented above is simply computed as:

1) The collective impact of the WINNN Programme (T) compared to the absence of the WINNN Programme (C), where terms are the means of the outcomes variable(s) of interest::

$$DD_{impact} = (T_1 - T_0) - (C_1 - C_0)$$

To supplement the raw difference-in-difference estimates, econometric models will be estimated to control for other factors that may co-determine the impact indicators. Given the selection process, it is likely that statistically significant differences will be observable between the two groups. Econometric models will be further refined in the first stage of the research project. In addition it may be possible to use routine and secondary sources to produce an interrupted time series analysis (ITS), which looks at changes to longer term trends (allowing us not only to compare changes from baseline levels between the groups, but also between the trends in indicators in those different areas, before and after the intervention). Secondary data sources that may be used

include the DHS, MICS and UNICEF SMART surveys (although only the SMART surveys can produce LGA-level data).

It is important to note that there will inevitably be issues of external validity i.e. the extent to which the impact results can be generalised (to a fully scaled-up programme or to other contexts and countries). Because the States and LGAs with each state covered by the impact evaluation were not chosen randomly means that the impact results will be representative of the WINNN Programme areas, but not representative of Northern Nigeria as a whole. To better understand issues of external validity, it will be important to fully appreciate the context in which the intervention is implemented, and indeed, the context of Northern Nigeria as a whole. To this end, the impact evaluation team plan to engage the implementing partners, relevant stakeholders and other components of ORIE (such as the operational research) to provide credible interpretations of the results, both for the WINNN Programme area, Northern Nigeria as a whole and for other contexts.

Implementation effectiveness

As described above, the quantitative impact evaluation will aim to assess the 'collective' impact of the WINNN Programme for those interventions concentrated in the focal LGAs. This means the impact of the programme in each State will be pooled together to arrive at an overall impact estimate for the entire programme. Because the WINNN Programme is being implemented in five States by two different implementing partners (see Table 2), it will be important to measure the implementation effectiveness of the Programme in each State to monitor any potential poor implementation to understand the impact estimate.

The results of the follow-up survey will be linked back to the level of implementation identified through the WINNN Programme M&E system in order to understand impacts, or indeed non-impacts. It is in this manner that the impact evaluation will also be able to distinguish between 'implementation failures' and 'theory failure', where assumptions of the causal linkages between input, activities, outputs and outcomes are incorrect. If possible, these indicators will be incorporated into an overall index of 'implementation effectiveness' which can later be used in conjunction with impact measures to assess the influence of implementation effectiveness outcomes and impacts.

Specific indicators of implementation effectiveness will be determined in collaboration with the WINNN Programme Chief of Party and M&E officer for each of the WINNN Programme components (Outputs 1-4) and will depend largely on data collected through the WINNN Programme's M&E system. These might include, for example:

- Output 1 (CMAM programme)
    - Number of CMAM programme sites set up and maintained
    - Number of children admitted to CMAM facilities for treatment
    - Number of children completing treatment at CMAM facilities
- Output 2 (IYCF interventions)
    - Number of active community promoters of IYCF practices

- Number of events in which information about IYCF practices is disseminated

Possible areas of impact and indicators

While the WINNN Programme Logframe will guide the impact evaluation, a specific Theory of Change Model will be developed for key impact indicators. This Theory of Change Model will be developed in collaboration with the Operations Research work stream as the model will also be able to signal important areas for operations research. The Theory of Change Model will help to elucidate the mechanism through which exposure to WINNN Programme interventions leads to final outcome measures. Table 5 provides a broad overview of the possible areas of impact and related indicators.

Table 5: Possible areas of impact and indicators

WINNN Programme Component	Hypothesis	Impact indicators
OUTPUT 1 (CMAM programme)	Timely detection of severe acute malnutrition (SAM) in a community and provision of treatment with RUTF for children without complications and combined with facility-based approaches for children with complications will reduce the prevalence of SAM	Under 5 prevalence of moderate and severe wasting
OUTPUT 2 (IYCF interventions)	Community-based approaches to deliver effective feeding practices including breastfeeding interventions will improve breastfeeding, weaning and feeding practices	Under 5 prevalence of wasting Under 5 prevalence of underweight Under 5 prevalence of stunting Proportion of infants 0-6 months exclusively breastfed Proportion of children 6-23 months receiving foods from four or more food groups supplementation in the last 6 months Proportion of women who took iron supplementation for at least 90 days during pregnancy for latest birth Early initiation of breastfeeding

It is important to note that while the WINNN Programme Logframe indicates that under five child mortality to be a key impact indicator, the quantitative impact evaluation will not measure this directly. Specialised surveys, such as the MICS or the DHS are required to accurately collect data on this indicator. As such, the impact evaluation will rely on secondary data (from the DHS and MICS) to report broad trends in under five child mortality in the WINNN Programmes States. It is uncertain if these sources of secondary data are able to provide estimates of this indicator at the level of the LGA.

Possible confounding or co-determinant variables

The determinants of malnutrition are multi-dimensional and inter-related. As such, the impact evaluation will collect data on more distal, underlying causes of malnutrition in order to control for determinants that are not directly related to the WINNN Programme package of interventions. The list of co-determinants is derived from the well-known UNICEF conceptual framework of child malnutrition and the ORIE Evidence Review.

Co-determinant variables that will also be included in the data collection are listed below.

Area of impact	Co-determinant indicators
Access to food	Dietary diversity Food security Access to land
Maternal and child care practices	Initiation of breastfeeding Appropriate complimentary feeding Mother attendance at antenatal care during pregnancy Immunisations Maternal health care seeking Duration of breastfeeding Hand washing practices
Water and sanitation services / environment	Availability of adequate sanitation Availability of safe water Environmental safety Availability of adequate shelter
Health services	Access to health care Quality of available health care
Socio-economic	Mother's education Mother's wealth Household wealth Mother's occupational status Main source of household livelihood Household size Woman's empowerment / decision making
Household member characteristics	Mother's parity Low birth weight of child Birth interval
Household characteristics	Dependency ratio

Time to follow-up

The time to follow-up will be three years from the baseline. This is to allow time for the CMAM and IYCF interventions to be implemented at long enough for it to have achieved sufficient coverage within an LGA to detect impact using a population based survey. Furthermore, health behaviour change interventions typically require longer periods of exposure before attitudes and practices begin to change. Finally, if a panel survey design is decided upon, a follow-up period beyond three years would risk high rates of attrition or high costs of tracking households for the follow-up survey.

Sampling strategy of the household survey

The baseline and follow-up survey (three years later) will form the core of the quantitative impact evaluation. It is expected that the baseline study will be conducted before the WINNN Programme intervention begins. Both surveys will include a household survey that will capture anthropometrics of all children under five years old.

In order to increase the precision of the impact evaluation and the likelihood of detecting an impact, the sample will be restricted to households that are most likely to be beneficiaries of the WINNN programme. These criteria include:

- Household with a woman of reproductive age
- Household with a child under five years old

Because children are most susceptible to the effects of such nutrition interventions between the ages of 0-24 months, we may consider over sampling households in the baseline survey with children under 12 months of age. As the impact evaluation will span three years, the Treatment Group should yield a sample of children that have been exposed to the WINNN Programme during the two years in which they are most vulnerable to malnutrition, and therefore the period of time in which we would most likely be able to detect the effects of the WINNN Programme.

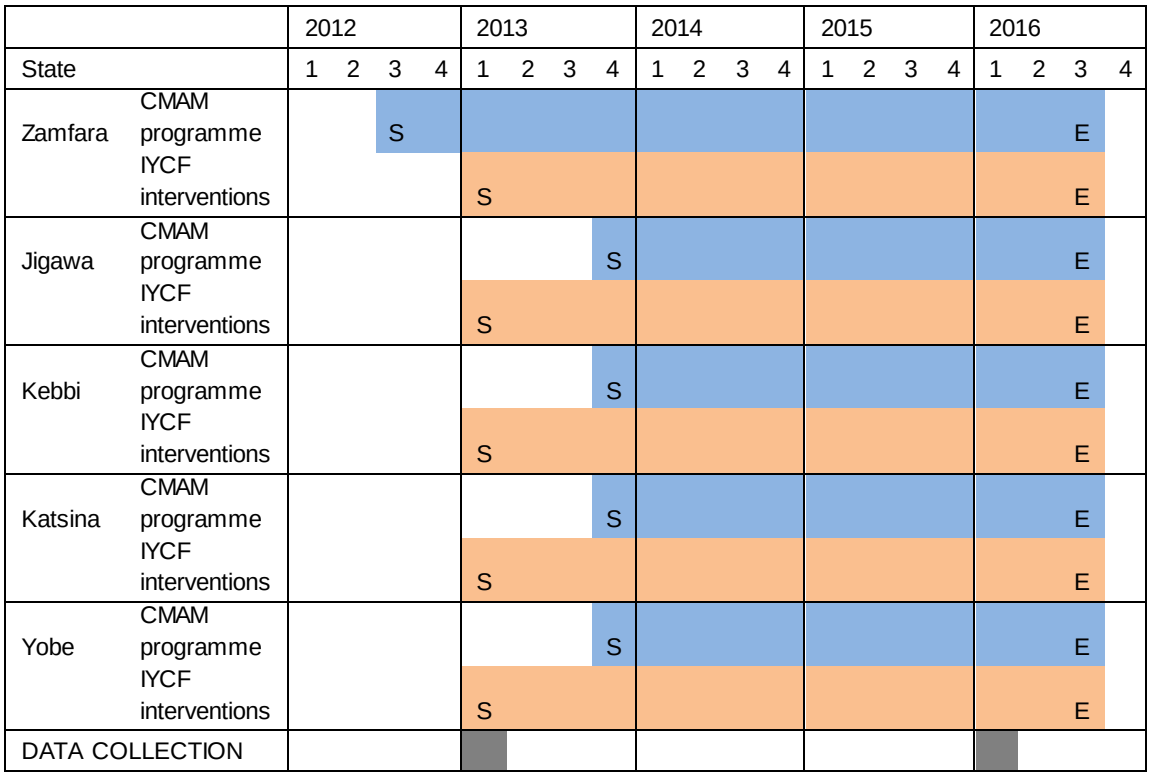
While over sampling households with children under 12 months will result in a cohort of children 'exposed' to WINNN Programme interventions during that key window in a child's life, it assumes that the WINNN Programme will be operating effectively with sufficient coverage soon after the baseline survey is conducted. The extent to which this assumption is a concern to the impact evaluation remains to be discussed within ORIE and the WINNN Programme.

It is important to note that the sampling strategy employed for the household survey will be representative of all households that meet the survey criteria (listed above) in the LGA as a whole as opposed to a sample of households from within the catchment area of the selected health facility. The rationale for this is because the WINNN Programme have committed to providing the CMAM and IYCF intervention to the entire LGA. Setting up the sampling strategy in this manner will allow the impact evaluation to assess the WINNN Programme's overall coverage and dimensions of geographic equity. Other dimensions of equity are discussed elsewhere.

Timing of the baseline and follow-up surveys

The roll-out of the WINNN Programme presents a particular challenge for the impact evaluation as WINNN has already begun operating some interventions in Jigawa and Zamfara and will expand into other States in subsequent years.

Because the WINNN Programme anticipates very similar roll-out for most States, the data collection for the impact evaluation could be done in two waves as suggested in the following figure where the baseline across all States is done in Quarter 1 of 2013 and the follow-up survey is done in Quarter 1 of 2016. Such a schedule will allow for three years of exposure to the WINNN intervention.



There are two complications that arise with this proposed schedule.

1) For Jigawa, Kebbi, Katsina and Yobe, the CMAM programme will begin 8-12 months after the start of the IYCF intervention in that State.

   This means that the baseline collected in these States may not be as accurate for the CMAM programme as no data will be collected immediately before it is implemented which is a problem only if the primary impact indicator for the CMAM programme, the prevalence of severe / moderate wasting, is particularly sensitive and likely to change between Quarter 1 2013 when the baseline survey is done and Quarter 4 2013 when the CMAM programme begins. The extent to which this indicator is sensitive to such changes will need to be discussed within ORIE and with the WINNN Programme.

   Conducting a 'second wave' of data collection or 'second baseline' just before the CMAM programme is launched in Quarter 4 2013 would have significant budgetary implications and require data to be collected in a season different to the first baseline which would limit the comparability of data between previous rounds of data collection.

2) The CMAM programme will already have been implemented in Zamfara for 6-9 months before the baseline survey is conducted.

   This is largely because the ORIE component of the WINNN Programme was contracted out much later than the implementation agencies of the WINNN Programme. The extent to which pre-baseline implementation of the CMAM programme will need to be carefully assessed. The CMAM programme targets severely / moderately wasted children under 5. According to the NDHS 2008, only 11% of children under five in Zamfara are wasted and 5% are severely wasted. Because the baseline will collect data on a sample of the general population, the extent to which 6-9 months of the CMAM programme intervention will have on general population levels of moderate / severe wasting may be negligent.

   Also, because of the phased roll out of the CMAM programme across Wards, by the time of the baseline survey, the WINNN Programme will only be implementing the CMAM programme in half of the total number of Wards that they are planning to cover within an LGA. At such an early stage in implementation, the impact on the population of children is likely to be minimal. Therefore, while the baseline in Zamfara may not be 'pure' (before any implementation), it should still be an effective baseline. To ensure the effectiveness of the baseline, a UNICEF SMART survey will be conducted in September 2012 to collect data on a limited number of key indicators which can then be checked against the 'late' baseline to ensure the assumption of minimal impact holds true.

Risks and limitations of the impact evaluation design

1) The time between baseline and follow-up is set at three years to allow the WINNN Programme to achieve adequate coverage and implementation effectiveness of their interventions to be able to detect a change in key impact indicators. If the WINNN Programme does not achieve effective implementation or adequate coverage during the evaluation period, impact on key population indicators may not be detected.

2) Related to (1) above, as the impact evaluation is designed to estimate the collective impact of the WINNN programme across all of the WINNN States, variable effectiveness of implementation between LGAs and States would reduce the measured impact of the WINNN Programme as whole. As discussed above, this will be monitored using secondary data largely from the WINNN M&E system.

3) The adequacy of non-WINNN LGA controls to provide a true counterfactual will need to be carefully assessed and monitored throughout the evaluation period. Should significant nutrition interventions be launched within control LGAs during the evaluation period, the impact of the WINNN risks under-estimation?

Security considerations

While the impact evaluation schedule outlined above reflects an optimal approach to the impact evaluation, security conditions in the WINNN State must be carefully considered before launching any activities.

Currently, Yobe State is known to be a highly dangerous and volatile State and launching any data collection activities in this State could represent significant dangers for the quantitative impact evaluation team. Therefore, it is recommended that this State be dropped from scope of the impact evaluation.

Quantitative research outputs and deliverables

High level work stream outputs include:

1. A baseline survey
2. A follow-up survey
3. A report on the collective impact of Output 1 and Output 2 of the WINNN Programme

Key deliverables

- Same as the high-level outputs listed above. Baseline and follow-up reports can be expected according to the schedule outlined above. The overall impact evaluation of the WINNN Programme can be expected 6 months after the collection of follow-up data in the final WINNN State (approximately June 2016)

International staffing plan for quantitative research

A competitive tender process has been launched to select the most capable survey organisation. After careful consideration of data collection proposals, one organisation will be selected and the in-country staffing plan will be developed with the selected partner

Quality assurance for quantitative research

A competitive tender process has been launched to select the most capable survey organisation. After careful consideration of data collection proposals, one organisation will be selected and quality assurance mechanisms of the selected partner, as well as OPM's survey quality assurance systems will be specified.

Annex C Sampling Strategy, sample size and power calculations

This annex outlines the sampling strategy used, the achieved sample size, the weights used, and the related power calculations for the present survey. Note that power calculations were done at baseline only. It also presents the response rate and item-non response analysis which gives more details about who was included in the final analysis, which is based on analysis done at endline.

C.1 Power Calculations

The purpose of this annex is to present estimates of the precision with which the quantitative impact evaluation is able to identify the impact of the WINNN interventions. It is important to emphasise that these are estimates that rely on a variety of assumptions explicitly laid out below. As previously explained, the aim of the quantitative impact evaluation is to estimate the effect of the WINNN intervention as a package in treatment LGAs versus no intervention in control LGAs. The following paragraphs will give an estimate of the size of the effect of the intervention that this study was identified to be able to detect at baseline.

Again, it is important to note that this analysis was implemented at baseline in order to say something about the effects that would potentially be identifiable with this study. Ex-post, i.e. at endline, there is no need to perform power calculations, given that effects have materialised and we can directly assess whether these are significantly different from zero or not, given the data at hand. These calculations are included here for completeness purposes.

Following standard results from theory on statistical testing, it is possible to identify, before the implementation of a survey, the sample size needed to test certain hypotheses on expected differences in means (or proportions) between two groups in a sample. In the present case, this could be to test whether the proportion of children malnourished in WINNN intervention areas before the intervention is statistically significantly different from the proportion after the WINNN intervention (SMART 2012b).

The needed sample size will depend on the difference in the values to be tested, the standard errors of the estimators, the required power of the test (i.e. the probability of correctly rejecting the hypothesis of no difference between values) and the required significance level of the test (i.e. the probability of falsely rejecting the hypothesis of no difference). Conversely, using a given sample size, a given estimate (mean or proportion) of an indicator, a required power of the test, and a significance level, it is possible to find the minimal difference to the given estimate that a statistical test will be able to identify (SMART, 2012b; Grosh and Munoz, 1996). In the context of a treatment versus control comparison, this is the minimum change in the outcome variable at which a statistically significant impact will be measured.

In addition to these standard procedures, clustered sampling needs to be taken into account in the present context. Because households and individuals within EAs are likely to have similar characteristics, and hence indicators will be correlated within these clusters, the standard errors of estimators will be larger than under simple random sampling (SMART, 2012b, p. 16 ff.). The factor by which standard errors using the clustered sampling method are inflated over standard errors using simple random sampling is called the Design Effect (DEFF), which for each indicator i is generally defined as follows:

$$(7) DEFF_i = 1 + (m - 1);$$

where m is the cluster size and ρ_i is the intra-class correlation (ICC) coefficient for indicator i, a measure of how much indicators are correlated with each other within clusters. When estimating the size of the detectable effect, inflated standard errors, and therefore the DEFF, need to be taken into account as well.

As can be seen, the size of the DEFF will generally depend on two factors: cluster size and the ICC. The formula above assumes constant cluster sizes. In the present context, however, cluster sizes vary. In some EAs, more children were interviewed than in others. In such instances, the DEFF should be defined differently so as to accommodate the varying cluster sizes.

There are several proposals in the literature setting out how this can be achieved, e.g. ESSEduNet (2013), Gabler et al. (2006), and Eldridge et al. (2006). We follow the approach suggested by Hemming et al. (2011), who recommend a procedure to adapt the DEFF to varying cluster sizes and who have developed a command to implement this procedure in Stata (Hemming and Marsh, 2013).

According to this approach, the DEFF with varying cluster sizes can be defined as follows:

$$(8)\ \mathrm{DEFF}_{\mathrm{i}}^{\mathrm{var}} = 1 + \{(\mathrm{cv}^2 + 1)\bar{\mathrm{m}} - 1\}\rho_{\mathrm{i}}$$

Here, cv is the coefficient of variation of cluster size, defined as the ratio of the standard deviation of cluster sizes to the mean cluster size, $\bar{\mathrm{m}}$ (Eldridge et al., 2006, p. 1293). The other terms are defined as before. When the sample size is known, the number of clusters is fixed but when the number of individuals might vary across clusters, which is the case in the present survey, the detectable difference, compared to simple random sampling, between two groups will be inflated by the square root of this DEFF (Hemming et al., 2011, p. 3).

For the following results we used the Stata command 'clustersampsi', developed by Hemming and Marsh (2013), in order to estimate the minimal detectable difference of indicators used in the ORIE survey for individuals in the treatment group – based on baseline data. For example, we wanted to estimate the minimal difference in stunting prevalence in treatment and control areas that we would be able to statistically detect after the implementation of the WINNN Programme.

For this exercise, we assumed that indicators would stay at baseline level in control areas, even after WINNN had been rolled out. In addition, we assumed that cluster sizes and ICCs would be the same. The ICC is estimated using the ANOVA estimator, implemented by 'l1way' in Stata, and allowing for varying cluster sizes and weights. The significance level is always set at 0.05% and power at 80%.

Throughout this study, clusters have been set at the level of the PSUs, i.e. the EA level, of which there were 420 in the treatment group at baseline. The average cluster size and the coefficient of variation of cluster size vary depending on the indicator analysed, and are hence presented below. For comparison purposes, the DEFF calculated using the approach outlined in Kish (1965), which is implemented using the Stata 'estat eff' command, is presented as well. Note also that for proportions of malnutrition the minimal detectable difference downwards, i.e. for a decreasing proportion, is reported. For the other indicators, the minimal difference detectable for an increasing outcome is reported in the table.

Table C 1 1. Design effects and minimum detectable differences for clusters at EA level

Indicator	Estimate	Average cluster size	Variation of cluster size	ICC at baseline	DEFF	Detectable difference	Alternative DEFF
Percentage stunted (0–35, LAZ/HAZ < -2SD)	52%	8	0.46	0.09	1.8	5 (58% to 53%)	1.9
Percentage underweight (0–35, WAZ < -2SD)	39%	8	0.45	0.13	2.1	5 (41% to 36%)	2.0
Percentage wasted (0–35, WLZ/WHZ < -2SD)	15%	7	0.46	0.02	1.2	3 (16% to 13%)	1.5
Exclusive breastfeeding among children aged < 6 months	9%	1	0.80	0.38	1.2	7 (7% to 14%)	1.8
Proportion of children with minimum dietary diversity (≥ 4 food groups) (aged 6–23 months)	15%	4	0.55	0.17	1.7	5 (14% to 19%)	1.7
Proportion of children receiving minimum acceptable diet (aged 6–23 months)	5%	4	0.55	0.11	1.5	3 (5% to 8%)	1.5
Percentage of children that are fully vaccinated for age (aged 12–23 months)	1%	3	0.61	0.19	1.6	3 (3% to 6%)	1.5

Notes: Estimates for the treatment areas only at baseline. Clusters are EAs.

Table C 1 1 above shows the sampling parameters for seven key indicators. Taking stunting for example, with an ICC of 0.09, an average cluster size of 8, a coefficient of variation of cluster size of about 0.5, and a DEFF of 1.8, the present sample will be sufficient to detect a decrease in stunting by four percentage points from 57% to 53%. Similarly, it will be sufficient to detect a decrease in underweight by four percentage points from 41% to 37% and by three percentage points in wasting from 15% to 12%. It will also be able to detect an increase in exclusive breastfeeding among children aged 0–5 months by eight percentage points (from 12% to 20%). The remaining estimates can be interpreted in an analogous way.

Clustering

As mentioned above, throughout this evaluation, and to present key impact estimates, standard errors were generally clustered at the level of EAs (PSUs). The rationale is that for this study, and for the purposes of comparing treatment and control areas, the report is simply presenting estimates for two separate populations for which the externally given LGAs are the universes.

However, in some instances, the impact evaluation literature suggests that, for the purposes of identifying causal effects, clusters should be defined at the level of the allocation of treatment. Note that, in section 3.3, we show that our key impact results are robust to such changes in the level of clustering. In the present case, these clusters were the 24 LGAs in the study. Again, the 12 treatment LGAs were purposefully chosen by state governments and WINNN IPs, so no random allocation of treatment was possible, and the number of treatment areas was fixed. Table C-9 below shows how the power estimates at baseline change when defining LGAs as clusters. As might be expected, because individuals within any LGA are more heterogeneous than within an EA, the ICC decreases. However, the DEFF for all indicators increases as well, due to increased average cluster size. This will increase standard errors and inflates the estimated minimal detectable difference. This means that estimates are less precise than identified above. In section 3.3, this is reflected in comparatively larger standard errors when clustering is set at the LGA level.

Table C 1 2. Design effects and minimum detectable differences with clustering at LGA level

Indicator	Estimate	Average cluster size	Variation of cluster size	ICC at baseline	DEFF	Detectable difference	Alternative DEFF
Percentage stunted (0–35, LAZ/HAZ < -2SD)	52%	266	0.17	0.05	14.6	14 (58% to 44%)	14.6
Percentage underweight (0–35, WAZ < -2SD)	39%	272	0.16	0.07	20.5	15 (41% to 26%)	18.7
Percentage wasted (0–35, WLZ/WHZ < -2SD)	15%	262	0.17	0.00	2.1	4 (16% to 12%)	2.1
Exclusive breastfeeding among children aged < 6 months	9%	47	0.26	0.07	4.4	12 (7% to 19%)	5.8
Proportion of children with minimum dietary diversity (≥ 4 food groups) (aged 6–23 months)	15%	138	0.19	0.06	9.5	12 (14% to 26%)	10.5
Proportion of children receiving minimum acceptable diet (aged 6–23 months)	5%	137	0.19	0.02	3.8	5 (5% to 10%)	3.7
Percentage of children that are fully vaccinated for age (aged 12–23 months)	1%	88	0.21	0.07	7.4	9 (3% to 12%)	7.6
Notes: Estimates for the treatment areas only at baseline. Clusters are LGAs.							

C.2 Response rates

This section presents response rate calculations for key observations included in the ORIE baseline and endline analysis.

Table C 1 Response rate analysis

Level of analysis	Baseline	Endline
Household		
Households interviewed	3,457	3,229
Households eligible (i.e. with children aged 0-35 months)	3,355	2,722
Response rate (households included in our analysis over households approached)	97.0%	84.3%
Children (0-35 months)		
Children (0-35 months)	6,833	5,567
Children (0-35 months) with some responses in the questionnaire	6,709	5,555
Response rate (Children with some questionnaire response over children included in our analysis)	98.2%	99.8%
Mothers (15-49 years)		
Mothers to children 0-35 months (mothers of reproductive age 15-49 years)	5,708	4,784
Mothers to children 0-35 months (mothers of reproductive age 15-49 years) with some responses in the questionnaire	5,670	4,765
Response rate (mother with some questionnaire response over mothers included in our analysis)	99.3%	99.6%

C.3 Item non-response

This annex presents the item-non response analysis for the baseline and endline indicators presented in Volume 1 and Volume 2. The following tables present item non-response rates for selected key indicators presented in this report. The base population for all indicators are the total number of individuals (of a certain age or sex) for which questionnaires were completed. The non-response is defined as one minus the quotient of observations included in the analysis over the base population.

For most indicators, the item non-response rate is below 2%. For indicators above this threshold at endline, a brief explanation is provided following the table. For indicators above this threshold at baseline, please refer to the baseline report (ORIE baseline report, 2013).

Table C3 1. IYCF counselling exposure

Indicator name	Baseline			Endline		
	Base population (n1)	Observations included in the analysis (n2)	Item non-response (1-n2/n1)	Base population (n1)	Observations included in the analysis (n2)	Item non-response (1-n2/n1)
Proportion of mothers (15-49 years) who have …						
… ever attended IYCF counselling in the community	5,708	5,666	0.7%	4,784	4,765	0.4%

Table C3 2. IYCF counselling exposure (endline only)

	Endline		
	Base population (n1)	Observations included in the analysis (n2)	Item non-response (1-n2/n1)
Proportion of mothers (15-49 years) who have …			
… attended ANC session and received IYCF counselling at any ANC session 1/	4,784	4,768	0.3%
…attended postnatal care and received IYCF counselling at any postnatal check up 1/	4,784	4,765	0.4%
… received IYCF counselling at the community and health facility 2/	4,784	4,765	0.4%
…received IYCF counselling at the community or at health facility 2/	4,784	4,765	0.4%
… ever heard about food demonstration sessions at community	4,784	4,765	0.4%
… ever attended food demonstration sessions at community	4,784	4,765	0.4%

Table C3 3. MNCH Exposure

Indicator name	Baseline			Endline		
	Base population (n1)	Observations included in the analysis (n2)	Item non-response (1-n2/n1)	Base population (n1)	Observations included in the analysis (n2)	Item non-response (1-n2/n1)
Proportion of mothers (15-49 years) who…:						
… have ever heard about MNCHW events	5,708	4,090	28.3%	4,784	4,765	0.4%

… have attended the last MNCHW events 1/	5,708	4,089	28.4%	4,784	4,765	0.4%

Table C3 4. CMAM programme exposure

Indicator name	Baseline			Endline		
	Base population (n1)	Observations included in the analysis (n2)	Item non-response (1-n2/n1)	Base population (n1)	Observations included in the analysis (n2)	Item non-response (1-n2/n1)
Proportion of children who have ever had their MUAC measured (children 6-35 months) 1/	5,686	5,686	0.0%	4,611	4,605	0.1%
Proportion of children 0-35 months who…:						
… were ever taken to any health centre for treatment through the CMAM programme /1	6,833	6,694	2.0%	5,567	5,554	0.2%
… were ever taken to an OTP facility for treatment with RUTF /1	6,833	6,694	2.0%	5,567	5,554	0.2%
… were ever taken to an ITP facility for treatment with F75/F100/RUTF /1	6,833	6,694	2.0%	5,567	5,554	0.2%

Table C3 5. Supplements

Indicator name	Baseline			Endline		
	Base population (n1)	Observations included in the analysis (n2)	Item non-response (1-n2/n1)	Base population (n1)	Observations included in the analysis (n2)	Item non-response (1-n2/n1)
Children 6-35 months who…:						
Received vitamin A in the last 6 months	5,686	5,651	0.6%	4,611	4,594	0.4%
Ever received vitamin A	5,686	5,673	0.2%	4,611	4,611	0.0%
Children 0-35 months who…:						
Received deworming medication in the last 6 months	6,828	6,626	3.0%	5,567	5,567	0.0%

Table C3 6. Supplements received by children (endline only)

Indicator name	Endline		
	Base population (n1)	Observations included in the analysis (n2)	Item non-response (1-n2/n1)
Children 0-35 months who…:			
Received ORS medication in the last 6 months	5,567	5,567	0.0%

Table C3 7. Supplements received during pregnancy (endline only)

Indicator name	Endline

	Base population (n1)	Observations included in the analysis (n2)	Item non-response (1-n2/n1)
During pregnancy, child's mother received/bought…(children 0-35 months):			
Iron supplements during pregnancy?	5,567	5,517	0.9%
Folic Acid supplements during pregnancy?	5,567	5,513	1.0%
Drugs for intestinal worms during pregnancy?	5,567	5,371	3.5%

- Identifying drugs for intestinal worms was difficult for mothers interviewed in this survey. For the purposes of this indicator, when the mother responded that she did not know whether she had received intestinal worms, the observations was set to missing. Therefore, the item non response on that indicator is relatively high.

Table C3 8. IYCF practices

Indicator name	Baseline			Endline		
	Base population (n1)	Observations included in the analysis (n2)	Item non-response (1-n2/n1)	Base population (n1)	Observations included in the analysis (n2)	Item non-response (1-n2/n1)
Logframe IYCF practice indicators:						
Early initiation to breastfeeding (<24h) breastfeeding (children 0-23 months)	4,452	4,402	0.2%	3,737	3,716	0.6%
Exclusive breastfeeding (children 0-5 months)	1,142	1,132	0.1%	949	948	0.1%
Minimum dietary diversity (children 6-23 months)	3,308	3,278	0.9%	2,783	2,783	0.0%
Other breastfeeding indicators:						
Ever breastfed (children 0-23 months)	4,452	4,409	0.1%	3,737	3,730	0.2%
Continued breastfeeding at one year (children 12-15 months)	913	913	0.0%	753	753	0.0%
Continued breastfeeding at two years (children 20-23 months)	644	644	0.0%	486	486	0.0%
Other complementary feeding indicators:						
Introduction to solid, semi-solid and soft foods (children 6-8 months)	571	571	0.0%	499	499	0.0%
Received at least two milk feedings during previous day (children 6-23 months who are not currently breasted)	658	658	0.0%	434	434	0.0%
Minimum meal frequency (children 6-23 months)	3,308	3,291	0.5%	2,783	2,767	0.6%
Minimum acceptable diet (children 6-23 months)	3,308	3,308	0.0%	2,783	2,782	0.0%

Table C3 9. IYCF practices (endline only)

Indicator name	Endline		
	Base population (n1)	Observations included in the analysis (n2)	Item non-response (1-n2/n1)

Logframe IYCF practice indicators			
Early initiation to breastfeeding (<1h) breastfeeding (children 0-23 months)	3,737	3,716	0.6%

Table C3 10. Knowledge of IYCF practices

Indicator name	Baseline			Endline		
	Base population (n1)	Observations included in the analysis (n2)	Item non-response (1-n2/n1)	Base population (n1)	Observations included in the analysis (n2)	Item non-response (1-n2/n1)
Proportion of mothers (15-49 years) who know that…:						
Colostrum is good for the baby and should be given to her/him	5,708	5,663	0.8%	4,765	4,765	0.0%
Water should not be given to children under 6 months	5,708	5,633	1.3%	4,784	4,765	0.4%
It is OK to feed a young baby under six months whenever he/she wants (non-standard feeding times)	5,708	5,653	1.0%	4,765	4,765	0.0%
Baby should only receive breastmilk for six months	5,708	5,662	0.8%	4,784	4,765	0.4%

Table C3 11. Vaccination

Indicator name	Baseline			Endline		
	Base population (n1)	Observations included in the analysis (n2)	Item non-response (1-n2/n1)	Base population (n1)	Observations included in the analysis (n2)	Item non-response (1-n2/n1)
Children 12- 35 months who received…:						
BCG vaccine	4,465	4,436	0.7%	3,634	3,575	1.6%
DPT/PENTA 1 vaccine	4,566	4,566	0.0%	3,634	3,634	0.0%
DPT/PENTA 2 vaccine	4,566	4,566	0.0%	3,634	3,634	0.0%
DPT/PENTA 3 vaccine	4,566	4,566	0.0%	3,634	3,634	0.0%
Polio 0 vaccine	4,465	4,431	0.8%	3,634	3,209	11.7%
Polio 1 vaccine	4,465	4,379	1.9%	3,634	3,625	0.3%
Polio 2 vaccine	4,465	4,379	1.9%	3,634	3,625	0.3%
Polio 3 vaccine	4,465	4,379	1.9%	3,634	3,625	0.3%
Measles vaccine	4,440	4,440	0.0%	3,634	3,516	3.2%
Fully immunized 1/	4,339	4,339	0.0%	3,634	3,487	4.0%

1/ According to DHS, full immunization requires: 1 shot of BCG vaccine, 1 shot of measles vaccine, 3 shots of polio (excluding at birth) and 3 shots of DPT/PENTA.

- Fully immunised item non-response is high because this indicator is missing if any of the vaccines included in this composite indicators are missing.

- Many children in Northern Nigeria do not have a vaccination card. For the purpose of the measles and polio 0 vaccines indicators, respondents were asked whether they got

measles vaccines if they do not have a vaccination card. Identifying vaccines is difficult for mothers interviewed in this survey-therefore many of them responded that they did not know whether their child was vaccinated against measles or polio 0 and the observations were set to missing. Therefore, the item non response on that indicator is relatively high.

Table C3 12. Malnutrition

Indicator name	Baseline				Endline			
	Base population (n1)	Observations included in the analysis (n2)	Observations not included in the analysis due to WHO outlier rule	Item non-response (1-n2/n1)	Base population (n1)	Observations included in the analysis (n2)	Observations not included in the analysis due to WHO outlier rule	Item non-response (1-n1/n2
Length for age z-score	6,833	6,410	255	6.2%	5,567	5,419	84	2.7%
Weight for age z-score	6,833	6,530	131	4.4%	5,567	5,472	34	1.7%
Weight for height z-score	6,833	6,278	249	8.1%	5,567	5,417	138	2.7%
Wasted (children 6-35 months)	5,686	5,306	96	6.7%	4,611	4,504	102	2.3%
Severely wasted (children 6-35 months)	5,686	5,306	96	6.7%	4,611	4,504	102	2.3%
Stunted (children 0-35 months)	6,833	6,410	255	6.2%	5,567	5,419	84	2.7%
Underweight (children 0-35 months)	6,833	6,530	131	4.4%	5,567	5,472	34	1.7%

- The item non-response rate for all anthropometric indicators is driven by extreme measurement biases that are filtered out of the anthropometric analysis as suggested by the WHO methodology. For length/height-for-age, observations with z-scores smaller than -6 or larger than 6 were dropped. For weight-for-height/length, observations with z-scores smaller than -5 or larger than 5 were dropped. Finally, for weight-for-age, observations with z-scores smaller than -6 or larger than 5 were dropped from the analysis. This fact increases item non-response above the level that would be expected by pure missing information. Note that the fact of lower levels of item non-response at endline compared to baseline, indicates that data quality improved between the two survey rounds. See section 3.2 for more detail.

Table C3 13. Malnutrition (endline only)

Malnutrition (endline)	Endline		
	Base population (n1)	Observations included in the analysis (n2)	Item non-response (1-n2/n1)
Children with SAM based on MUAC or oedema (children 6-35 months)	4,543	4,504	0.0%
Children with SAM based on MUAC, oedema and WHZ (children 6-35 months)	4,611	4,486	2.7%

C.4 Sampling Weights

In order to obtain estimates of key indicators that are representative for WINNN intervention areas, the observed values were analysed using sampling weights that were equal to the inverse of the probabilities of the observations to be selected into the sample.

EAs

In analyses at the community level, the first sampling stage, this meant that observations were weighted by the inverse of the probability of an EA to be selected within a certain LGA:

$$(9)\ \mathrm{p}_{\mathrm{i}}^{\mathrm{EA}} = \frac{35}{\mathrm{N}_{\mathrm{i}}^{\mathrm{EA}}}$$

where 35 is the total number of EAs to be selected in each LGA, $N_{i^{EA}}$ is the total number of EAs listed in LGA i, and $p_{i^{EA}}$ is hence the probability of selection in LGA i. Analyses at the community level were implemented using appropriately normalised values of weights derived from these probabilities.

HHs

At the household level – i.e. the second sampling stage – the probability of selection was given as follows:

$$(10)\ \ \mathrm{p}_{\mathrm{ij}}^{\mathrm{HH}} = \mathrm{p}_{\mathrm{i}}^{\mathrm{EA}} \times \frac{4}{\mathrm{N}_{\mathrm{j}}^{\mathrm{HH}}}$$

where p_{iEA} is defined as above, 4 is the total number of households to be selected within each EA, $N_{j^{HH}}$ is the total number of households listed in EA j, and $p_{ij^{HH}}$ is therefore the probability of selection of the four households in EA j and LGA i.

Analyses at household and individual level (mothers and children) were implemented using appropriately normalised inverse values of these probabilities as weights. Note that each individual within a household had the household's probability of being selected, as all children and all women of the relevant age were interviewed in each household.

Attrition correction

Note that for the purposes of the final analysis presented in Volume 1, weights were corrected to take attrition into account. Please see section 3.1.3 on the technical detail of how this was implemented.

C.5 Precision of estimates

The purpose of this annex is to present confidence intervals for key results presented in Volume 1 and Volume 2, and hence to give an idea for the precision of these estimates, taking into account sampling error.

Table C 5 1. IYCF counselling exposure

Indicator name	Treatment			Control		
	Baseline	Endline	Diff (EL-BL)	Baseline	Endline	Diff (EL-BL)
Proportion of mothers (15-49 years) who ever attended IYCF counselling in the community	7.5	31.5	24.0***	4.3	7.4	3.1**
95% confidence interval	(6.3,9.0)	(27.9,35.3)		(3.4,5.5)	(5.6,9.7)	
N	*2,833*	*2,303*		*2,833*	*2,235*	

Table C 5 2. IYCF counselling exposure (endline only)

Indicator name	Endline			
	Total	Treatment	Control	Diff (T-C)
Proportion of mothers (15-49 years) who have …				
… attended ANC session and received IYCF counselling at any ANC session	35.3	42.2	28.5	13.7***
95% confidence interval	(32.7,37.9)	(38.2,46.3)	(25.4,31.8)	
N	*4,541*	*2,305*	*2,236*	
…attended postnatal care and received IYCF counselling at any postnatal check up	16.6	20.2	12.9	7.3***
95% confidence interval	(14.8,18.5)	(17.5,23.3)	(10.9,15.4)	
N	*4,538*	*2,301*	*2,237*	
… received IYCF counselling at the community and health facility	11.9	19.7	4.2	15.4***
95% confidence interval	(10.3,13.6)	(16.9,22.7)	(3,5.9)	
N	*4,556*	*2,314*	*2,242*	
…received IYCF counselling at the community or at health facility	47.1	58.3	36.1	22.2***
95% confidence interval	(44.5,49.8)	(54.2,62.4)	(32.7,39.6)	
N	*4,556*	*2,314*	*2,242*	
… ever heard about food demonstration sessions at community	17.3	26.0	8.8	17.2***
95% confidence interval	(15.5,19.3)	(22.9,29.5)	(7.2,10.8)	
N	*4,538*	*2,303*	*2,235*	
… ever attended food demonstration sessions at community	6.0	11.0	1.1	9.8***
95% confidence interval	(4.9,7.4)	(8.8,13.5)	(0.5,2.4)	
N	*4,538*	*2,303*	*2,235*	
Mean number of times mothers have received IYCF community counselling in last 6 months	1.0	1.1	0.6	0.5***
95% confidence interval	(0.9,1.1)	(0.9,1.2)	(0.4,0.8)	
N	*877*	*732*	*145*	
Mean number of people who participated in group counselling last time	30.5	30.0	32.5	-2.5
95% confidence interval	(27,33.9)	(26.3,33.7)	(23.1,41.9)	
N	*526*	*440*	*86*	

Table C 5 3. IYCF counselling setting (endline only)

Indicator name	Endline			
	Total	Treatment	Control	Diff (T-C)
Proportion of mothers (15-49 years) who attended training on IYCF practices at….				
Own house or neighbour's house	56.2	56.7	54.0	2.7
95% confidence interval	(50.5,61.6)	(50.5,62.6)	(40.0,67.3)	
N	*873*	*728*	*145*	
Community volunteer's house	6.4	7.5	2.0	5.5**
95% confidence interval	(4.3,9.3)	(5.0,11)	(0.6,6.6)	
N	*873*	*728*	*145*	
Family or community ceremony	5.9	6.7	2.8	3.8

95% confidence interval	(3.8,9.0)	(4.2,10.5)	(1.1,7.3)	
N	*873*	*728*	*145*	
Village head's house	28.1	27.9	28.8	-0.9
95% confidence interval	(23.3,33.4)	(22.7,33.9)	(18.4,42)	
N	*873*	*728*	*145*	
Public space in the community	16.0	15.6	18.0	-2.4
95% confidence interval	(12.5,20.4)	(11.7,20.5)	(10.6,28.8)	
N	*873*	*728*	*145*	
Proportion of mothers (15-49 years) who attended training on IYCF practices led by….				
Community Volunteer	70.0	71.6	63.1	8.5
95% confidence interval	(65.1,74.4)	(66.3,76.4)	(50.2,74.3)	
N	*883*	*738*	*145*	
Medical / health facility staff	27.3	28.3	23.3	5.0
95% confidence interval	(22.5,32.7)	(22.7,34.5)	(14.9,34.5)	
N	*883*	*738*	*145*	
Someone else	16.9	16.6	18.1	-1.5
95% confidence interval	(13.2,21.3)	(12.4,21.8)	(11.3,27.6)	
N	*883*	*738*	*145*	
Do not know	1.8	1.6	2.9	-1.3
95% confidence interval	(0.9,3.5)	(0.7,3.3)	(0.7,10.8)	
N	*883*	*738*	*145*	
Proportion of mothers (15-49 years) who attended IYCF…:				
One-to-one counselling	25.8	25.1	28.7	-3.6
95% confidence interval	(21.7,30.4)	(20.7,30.1)	(19.2,40.5)	
N	*883*	*738*	*145*	
Group counselling	70.9	71.0	70.5	0.5
95% confidence interval	(66.4,75)	(66.1,75.5)	(58.8,80)	
N	*883*	*738*	*145*	
Both counselling	3.3	3.8	0.8	3.0*
95% confidence interval	(1.9,5.7)	(2.2,6.8)	(.1,5.9)	
N	*883*	*738*	*145*	

Table C 5 4. MNCHW events exposure

Indicator name	Treatment			Control		
	Baseline	Endline	Diff (EL-BL)	Baseline	Endline	Diff (EL-BL)
Proportion of mothers (15-49 years) who …						
… have ever heard about MNCHW events	13.2	42.9	29.8***	10.5	35	24.5***
95% confidence interval	(11.2,15.4)	(39.1,46.9)		(8.7,12.7)	(31.3,39.0)	
N	*2,010*	*2,303*		*2,080*	*2,235*	
… have attended the last MNCHW events	5.0	12.9	7.9***	4.8	7.2	2.4*
95% confidence interval	(3.7,6.8)	(10.5,15.8)		(3.5,6.6)	(5.4,9.4)	
N	*2,009*	*2,303*		*2,080*	*2,235*	

Table C 5 5. MNCH services

Indicator name	Treatment			Control		
	Baseline	Endline	Diff (EL-BL)	Baseline	Endline	Diff (EL-BL)
Proportion of mothers (15-49 years who went to the last MNCHW event) whose children received Vitamin A drops	93.0	76.6	-16.3***	92.7	78.8	-13.9*
95% confidence interval	(85.8,96.6)	(69.6,82.5)		(84.9,96.7)	(68.6,86.3)	
N	*98*	*297*		*88*	*151*	

Table C 5 6. MNCH services (endline only)

	Endline			
	Total	Treatment	Control	Diff (T-C)
Proportion of mothers (15-49 years who went to the last MNCHW events) whose children received the following:				
Deworming pills	61.4	61.1	62.0	-0.9
95% confidence interval	(54.5,67.8)	(52.4,69.1)	(50.5,72.2)	
N	*438*	*292*	*146*	
Malnutrition examination with MUAC	42.1	45.0	36.9	8.1
95% confidence interval	(36.7,47.7)	(38.2,52.1)	(28.5,46.2)	
N	*444*	*296*	*148*	
ORS	66.3	70.5	58.4	12.1
95% confidence interval	(49.1,80.1)	(47.9,86.1)	(35.8,78)	
N	*58*	*40*	*18*	
Zinc	65.8	66.9	63.8	3.1
95% confidence interval	(48.8,79.5)	(45.9,82.8)	(34.8,85.4)	
N	*53*	*36*	*17*	
Proportion of mothers (15-49 years who went to the last MNCHW event) who received :				
Long Lasting Insectidal Nets (LLINs)	51.2	52.8	48.5	4.3
95% confidence interval	(44.6,57.7)	(44.2,61.2)	(38.4,58.6)	
N	*450*	*299*	*151*	
Tetanus toxoid vaccine	39.8	38.7	41.7	-3.0
95% confidence interval	(34.1,45.7)	(31.8,46.1)	(32.2,51.8)	
N	*441*	*292*	*149*	
Proportion of pregnant women (15-49 years who went to the last MNCHW events) who received :				
Folic Acid/Iron Folate	73.8	73.9	73.6	0.4
95% confidence interval	(57.6,85.4)	(55.4,86.6)	(42.9,91.2)	
N	95	59	36	
Iron supplements	80.0	83.8	74.7	9.1
95% confidence interval	(65,89.6)	(71.1,91.6)	(43.4,91.9)	
N	95	59	36	

Table C 5 7. Reason for not attending MNCHW events

Indicator name	Treatment			Control		
	Baseline	Endline	Diff (EL-BL)	Baseline	Endline	Diff (EL-BL)
Reason why mothers (15-49 years) did not attend MNCHW events:						
Have not ever heard of MNCHW	91.6	65.5	-26***	94	70	-24.1***

95% confidence interval	(89.6,93.2)	(61.9,69.0)		(92.5,95.3)	(66.2,73.5)	
N	*1,908*	*2,004*		*1,987*	*2,084*	
Did not have time	3.4	6.8	3.4***	1.7	6	4.3***
95% confidence interval	(2.5,4.5)	(5.3,8.7)		(1.1,2.6)	(4.8,7.6)	
N	*1,908*	*2,004*		*1,987*	*2,084*	
Too far / too expensive	0.8	2.0	1.1*	0.5	1.7	1.2**
95% confidence interval	(0.3,1.9)	(1.3,3.0)		(0.3,0.9)	(1.1,2.6)	
N	*1,908*	*2,004*		*1,987*	*2,084*	
Not useful	0.3	2.1	1.8***	0.2	2.4	2.2***
95% confidence interval	(0.1,0.6)	(1.5,2.9)		(0.1,0.6)	(1.6,3.5)	
N	*1,908*	*2,004*		*1,987*	*2,084*	
Did not have permission to go	1.2	7.8	6.7***	1.1	8	7.0***
95% confidence interval	(0.7,1.9)	(6.4,9.5)		(0.7,1.6)	(6.4,10)	
N	*1,908*	*2,004*		*1,987*	*2,084*	
Did not know about last MNCHW event	2.1	11.3	9.3***	2.0	9.5	7.4***
95% confidence interval	(1.4,3.0)	(9.5,13.5)		(1.3,3.3)	(7.6,11.7)	
N	*1,908*	*2,004*		*1,987*	*2,084*	
Other	0.7	5.7	5***	0.4	4.4	3.9***
95% confidence interval	(0.4,1.5)	(4.4,7.4)		(0.2,0.9)	(3.4,5.6)	
N	*1,908*	*2,004*		*1,987*	*2,084*	

Table C 5 8. CMAM programme exposure

Indicator name	Treatment			Control		
	Baseline	Endline	Diff (EL-BL)	Baseline	Endline	Diff (EL-BL)
Proportion of children who have ever had their MUAC measured (children 6-35 months)	12.9	20.1	7.1***	7.6	7.8	0.3
95% confidence interval	(11.4,14.7)	(17.9,22.4)		(6.3,9.1)	(6.4,9.5)	
N	*2,875*	*2,215*		*2,811*	*2,175*	
… were ever taken to any health centre for treatment through the CMAM programme	6.7	7.7	0.9	3.7	3.1	-0.6
95% confidence interval	(5.3,8.4)	(6.4,9.1)		(2.9,4.7)	(2.4,4.0)	
N	*3,383*	*2,669*		*3,312*	*2,621*	
… were ever taken to an OTP facility for treatment with RUTF	5.8	7.3	1.5	3.2	3.0	-0.2
95% confidence interval	(4.5,7.5)	(6.1,8.8)		(2.5,4.1)	(2.3,3.9)	
N	*3,382*	*2,669*		*3,312*	*2,621*	
… were ever taken to an SC facility for treatment with F75/F100/RUTF	0.9	0.4	-.5*	0.5	0.2	-0.3
95% confidence interval	(0.6,1.3)	(0.2,.8)		(0.3,.9)	(0.1,.5)	
N	*3,382*	*2,669*		*3,312*	*2,621*	

Table C 5 9. CMAM programme exposure (endline only)

Indicator name	Endline			
	Total	Treatment	Control	Diff (T-C)
Proportion of children with SAM who …				
… have ever had their MUAC measured (children 6-35 months)	23	29.6	15.1	14.5*
95% confidence interval	(17.6,29.5)	(21.0,39.9)	(9.8,22.5)	

N	*299*	*157*	*142*	
… were ever taken to any health centre for treatment through the CMAM programme (children 6-35 months)	14.7	17.5	11.3	6.2
95% confidence interval	(10.7,19.8)	(11.6,25.6)	(6.8,18.0)	
N	*299*	*157*	*142*	
Of children 6-35 months who ever had their MUAC measured:				
Proportion of children who had their MUAC measured in last 30 days	23.7	21.8	28.6	-6.9
95% confidence interval	(19.4,28.7)	(17.1,27.2)	(19.8,39.4)	
N	*596*	*415*	*181*	
Proportion of children had their MUAC measured in last 6 months	56.4	57.6	53.2	4.5
95% confidence interval	(50.9,61.6)	(51.1,63.9)	(43.8,62.3)	
N	*596*	*415*	*181*	
Of children 6-35 months who were examined for malnutrition in the last six months using MUAC:				
Number of times the child was examined in the last 6 months	2.8	2.8	2.8	0.0
95% confidence interval	(2.4,3.2)	(2.4,3.2)	(1.7,3.9)	
N	*334*	*232*	*102*	

Table C 5 10. Mean number of times children went to the OTP health facility last time they received treatment with RUTF

Indicator name	Endline			
	Total	Treatment	Control	Diff (T-C)
Mean number of times children went to the OTP health facility last time they received treatment with RUTF	5.3	5.4	5.1	0.3
95% confidence interval	(4.9,5.7)	(5,5.8)	(4.5,5.7)	
N	*275*	*190*	*85*	

Table C 5 11. Location of MUAC measurement

Indicator name	Endline			
	Total	Treatment	Control	Diff (T-C)
Proportion of children 6-35 months who had their MUAC measured at…:				
House	7.7	9.0	4.3	4.8
95% confidence interval	(5.3,10.9)	(6,13.3)	(1.9,9.3)	
N	*597*	*415*	*182*	
Community	8.3	7.2	11.1	-3.8
95% confidence interval	(5.8,11.8)	(4.6,11.2)	(6,19.5)	
N	*597*	*415*	*182*	
Health Facility	84.0	83.7	84.7	-0.9
95% confidence interval	(79.7,87.5)	(78.6,87.9)	(75.8,90.7)	
N	*597*	*415*	*182*	

Table C 5 12. RUTF Knowledge

Indicator name	Treatment			Control		
	Baseline	Endline	Diff (EL-BL)	Baseline	Endline	Diff (EL-BL)
Does RUTF need preparation before it can be fed to child? (No)	66.4	91.9	25.5***	68.6	94.3	25.7***
95% confidence interval	(52.5,77.9)	(85.1,95.8)		(57.7,77.8)	(87.0,97.6)	

N	*178*	*192*		*108*	*86*	

Table C 5 13. RUTF knowledge (endline only)

Indicator name	Endline			
	Total	Treatment	Control	Diff (T-C)
Proportion of carers who have ever taken their child to an OTP facility for treatment and who respond to:				
Shared the RUTF sachet? (No)	68.5	67.7	70.6	-2.9
95% confidence interval	(61.4,74.9)	(58.4,75.8)	(60.3,79.1)	
N	*278*	*192*	*86*	
Can you but RUTF sachets outside the health facility? (Yes)	16.7	17.8	14.2	3.7
95% confidence interval	(10.9,24.8)	(10.5,28.6)	(7.3,25.7)	
N	*278*	*192*	*86*	

Table C 5 14. IYCF practices

Indicator name	Treatment			Control		
	Baseline	Endline	Diff (EL-BL)	Baseline	Endline	Diff (EL-BL)
Early initiation (< 24 hours) to breastfeeding (children 0-23 months)	64.4	82.8	18.4***	60.2	72.9	12.7***
95% confidence interval	(60.9,67.7)	(79.9,85.3)		(56.5,63.8)	(69.3,76.2)	
N	*2,190*	*1,784*		*2,212*	*1,754*	
Exclusive breastfeeding (children 0-23 months)	9.2	19.5	10.3***	3.1	7.2	4.1*
95% confidence interval	(6.5,12.7)	(15.4,24.4)		(1.6,5.8)	(4.7,10.7)	
N	*578*	*453*		*554*	*446*	
Minimum dietary diversity (children 6-23 months)	14.5	20.8	6.3***	12.8	20.7	7.8***
95% confidence interval	(12.3,17.1)	(17.9,24.1)		(11,14.9)	(18,23.7)	
N	*1,616*	*1,339*		*1,662*	*1,315*	
Child ever breastfed (children 0-23 months)	99.6	99.9	.3*	99.7	99.8	0.1
95% confidence interval	(99.2,99.8)	(99.7,100)		(99.4,99.9)	(99.6,99.9)	
N	*3,392*	*2,669*		*3,314*	*2,622*	
Continued breastfeeding at one year (children 12-15 months)	90	97.6	7.6***	94.9	98.2	3.3*
95% confidence interval	(86.3,92.8)	(93.7,99.1)		(91.9,96.8)	(95.7,99.2)	
N	*454*	*354*		*459*	*373*	
Continued breastfeeding at two years (children 20-23 months)	27.2	29.8	2.6	28.7	33.7	5
95% confidence interval	(21.3,34.1)	(23,37.6)		(23.1,35)	(26.5,41.7)	
N	*308*	*232*		*336*	*230*	
Introduction to solid, semi-solid and soft foods (children 6-8 months)	73.6	68.7	-5.0	71.9	62.5	-9.4
95% confidence interval	(67.6,78.9)	(61.9,74.7)		(64.6,78.1)	(54.5,69.8)	
N	*287*	*246*		*284*	*222*	
Received at least two milk feedings during previous day (children 6-23 months who are not currently breasted)[†]	9.5	6.1	-3.4	10.5	3.2	-7.3**
95% confidence interval	(6,14.8)	(3.1,11.8)		(7.3,15)	(1.4,7.3)	
N	*312*	*205*		*346*	*215*	
Minimum meal frequency (children 6–23 months)	23.7	31.6	7.9***	21.5	25.4	3.9*

95% confidence interval	(21.2,26.5)	(27.9,35.5)		(19.3,24)	(22.4,28.7)	
N	*1,625*	*1,332*		*1,666*	*1,309*	
Minimum acceptable diet (children 6–23 months)	4.9	9.0	4.2***	4.5	9.0	4.6***
95% confidence interval	(3.7,6.4)	(7.1,11.4)		(3.4,5.8)	(7.4,11)	
N	*1,632*	*1,338*		*1,676*	*1,315*	

Table C 5 15. IYCF practices (endline only)

Indicator name	Endline			
	Total	Treatment	Control	Diff (T-C)
Early initiation (< 1 hour) to breastfeeding (children 0-23 months)	30.3	38.0	22.6	15.4***
95% confidence interval	(27.5,33.2)	(33.6,42.6)	(19.7,25.8)	
N	*3,538*	*1,784*	*1,754*	

Table C 5 16. IYCF practices knowledge

Indicator name	Treatment			Control		
	Baseline	Endline	Diff (EL-BL)	Baseline	Endline	Diff (EL-BL)
Colostrum should be given to baby	78.4	77.5	-0.9	73.5	74.9	1.5
95% confidence interval	(75.9,80.7)	(74.3,80.4)		(70.4,76.3)	(71.6,78)	
N	*2,831*	*2,303*		*2,832*	*2,235*	
Baby should receive only breastmilk for 6 months	36.7	45.1	8.4***	15.0	25.2	10.1***
95% confidence interval	(33.5,40.1)	(41.9,48.4)		(13.1,17.3)	(22.6,28)	
N	*2,831*	*2,303*		*2,831*	*2,235*	
Water should not be given to children under 6 months	8.3	19.1	10.8***	6.5	6.8	0.3
95% confidence interval	(6.8,10.1)	(16.5,22.1)		(5.2,8.1)	(5.3,8.7)	
N	*2,811*	*2,303*		*2,822*	*2,235*	
Baby should be fed whenever he/she wants	89.2	94.3	5.1***	95.4	96.1	0.7
95% confidence interval	(87.5,90.7)	(92.6,95.6)		(94.1,96.3)	(94.8,97.1)	
N	*2,824*	*2,303*		*2,829*	*2,235*	

Table C 5 17. IYCF practices knowledge (endline only)

Indicator name	Endline			
	Total	Treatment	Control	Diff (T-C)
Proportion of mothers (15-49 years) who knew that:				
Baby should start breastfeeding immediately or within the first hour	37.7	44.8	30.8	14***
95% confidence interval	(35.1,40.5)	(40.7,49)	(27.7,34.1)	
N	*4,529*	*2,298*	*2,231*	
Holy water (zamzam) should not be given to children under 6 months	10.9	14.6	7.3	7.2***
95% confidence interval	(9.6,12.4)	(12.5,16.9)	(5.8,9.2)	
N	*4,538*	*2,303*	*2,235*	

Table C 5 18. Micronutrient supplementation

Indicator name	Treatment	Control

	Baseline	Endline	Diff (EL-BL)	Baseline	Endline	Diff (EL-BL)
Ever received vitamin A (children 6-35 months)	45.3	43.2	-2.2	51.3	32.1	-19.2***
95% confidence interval	(42.2,48.5)	(39.9,46.5)		(48,54.7)	(28.8,35.6)	
N	*2,869*	*2,218*		*2,804*	*2,177*	
Received vitamin A in the last 6 months (children 6-35 months)	37.6	28	-9.6***	43.5	17.8	-25.6***
95% confidence interval	(34.7,40.7)	(25.1,31.1)		(40.4,46.6)	(15.4,20.6)	
N	*2,858*	*2,212*		*2,793*	*2,166*	
Has child received deworming medication in the last 6 months? (children 0-35 months)	8.2	11.5	3.3**	7	6.5	-0.5
95% confidence interval	(6.9,9.7)	(9.7,13.7)		(5.9,8.2)	(5.1,8.1)	
N	*3,353*	*2,677*		*3,273*	*2,625*	

Table C 5 19. Micronutrient supplementation (endline only)

Indicator name	Endline			
	Total	Treatment	Control	Diff (T-C)
Has child been given ORS in the last 6 months? (children 0-35 months)	21.0	23.0	19.0	4.0*
95% confidence interval	(19.4,22.6)	(20.8,25.3)	(16.9,21.3)	
N	*5,302*	*2,677*	*2,625*	

Table C 5 20. Micronutrient supplementation during pregnancy

Indicator name	Endline			
	Total	Treatment	Control	Diff (T-C)
During pregnancy, did the child's mother receive/buy…?				
Iron supplements (children 0-35 months)	57.4	62.3	52.5	9.8***
95% confidence interval	(54.8,59.9)	(58.8,65.7)	(48.8,56.2)	
N	*5,254*	*2,652*	*2,602*	
Folic Acid supplements (children 0-35 months)	53.6	58.4	48.9	9.5***
95% confidence interval	(51.1,56)	(55.2,61.6)	(45.4,52.5)	
N	*5,249*	*2,647*	*2,602*	
Drugs for intestinal worms during pregnancy (children 0-35 months)	23.8	28.7	19	9.8***
95% confidence interval	(21.8,25.8)	(25.7,32)	(16.5,21.7)	
N	*5,119*	*2,584*	*2,535*	

Table C 5 21. Vaccination

Indicator name	Treatment			Control		
	Baseline	Endline	Diff (EL-BL)	Baseline	Endline	Diff (EL-BL)
BCG vaccine	25.9	28.4	2.5	17.7	21	3.3*
95% confidence interval	(22.9,29.1)	(24.9,32.1)		(15.5,20.2)	(18.2,24.2)	
N	*2,247*	*1,702*		*2,189*	*1,711*	
Polio 0	17.4	38.2	20.7***	15.8	31.1	15.3***
95% confidence interval	(15.1,20.1)	(34.6,41.9)		(13.9,17.9)	(28.1,34.1)	

N	*2,250*	*1,558*		*2,181*	*1,505*	
Polio 1	55	75.7	20.7***	54.6	80.3	25.7***
95% confidence interval	(51.9,58)	(71.8,79.1)		(51.5,57.6)	(77.1,83.1)	
N	*2,210*	*1,722*		*2,169*	*1,738*	
Polio 2	53.2	72.7	19.6***	52.3	77.8	25.5***
95% confidence interval	(50.1,56.2)	(68.8,76.4)		(49.4,55.3)	(74.6,80.8)	
N	*2,210*	*1,722*		*2,169*	*1,738*	
Polio 3	51.1	71	19.9***	49.7	75.5	25.8***
95% confidence interval	(48.1,54.1)	(67,74.6)		(46.8,52.7)	(72.3,78.5)	
N	*2,210*	*1,722*		*2,169*	*1,738*	
DPT/PENTA 1 vaccine	19.5	8.7	-10.9***	13.4	7.5	-5.9***
95% confidence interval	(17.0,22.3)	(6.9,10.8)		(11.5,15.6)	(5.9,9.5)	
N	*2,315*	*1,729*		*2,251*	*1,739*	
DPT/PENTA 2 vaccine	3.1	6.7	3.6***	3.5	6.2	2.7***
95% confidence interval	(2.2,4.3)	(5.1,8.7)		(2.5,4.8)	(4.7,8.1)	
N	*2,315*	*1,729*		*2,251*	*1,739*	
DPT/PENTA 3 vaccine	2.2	5.2	3.1***	2.4	4.7	2.3***
95% confidence interval	(1.4,3.3)	(3.9,7)		(1.6,3.4)	(3.5,6.3)	
N	*2,315*	*1,729*		*2,251*	*1,739*	
Measles vaccine	16.8	33.2	16.4***	15.9	26.7	10.8***
95% confidence interval	(14.7,19.3)	(29.7,36.9)		(13.8,18.3)	(23.6,29.9)	
N	*2,248*	*1,663*		*2,192*	*1,695*	
Fully immunized	1	3.2	2.2***	1.5	3	1.4**
95% confidence interval	(0.5,1.7)	(2.2,4.6)		(1.0,2.3)	(2.1,4.2)	
N	*2,191*	*1,650*		*2,148*	*1,681*	

Table C 5 22. Malnutrition

Indicator name	Treatment			Control		
	Baseline	Endline	Diff (EL-BL)	Baseline	Endline	Diff (EL-BL)
Wasted (children 6-35 months)	14.9	17.6	2.7	17.5	17	-0.5
95% confidence interval	(13.1,16.9)	(15.7,19.6)		(15.9,19.3)	(15.1,19.2)	
N	*2,726*	*2,174*		*2,580*	*2,120*	
Severely wasted (children 6-35 months)	5.3	6.3	1.0	6.3	5.3	-1.0
95% confidence interval	(4.2,6.6)	(5.2,7.6)		(5.3,7.5)	(4.3,6.6)	
N	*2,726*	*2,174*		*2,580*	*2,120*	
Stunted (children 0-35 months)	52.1	49.2	-3.0	54.5	52.6	-1.9
95% confidence interval	(49.6,54.7)	(46.6,51.8)		(52.1,56.8)	(50.3,55)	
N	*3,306*	*2,606*		*3,104*	*2,554*	
Underweight (children 0-35 months)	38.7	38.5	-0.2	39.1	37.9	-1.2
95% confidence interval	(36.4,41)	(36.1,40.8)		(36.7,41.4)	(35.3,40.4)	
N	*3,329*	*2,641*		*3,201*	*2,571*	

Table C 5 23. Malnutrition (endline only)

Indicator name	Endline

	Total	Treatment	Control	Diff (T-C)
Proportion of children (6-35 months) who are:				
Severely acutely malnourished based on MUAC measurement and/or oedema	5.5	5.6	5.5	0.1
95% confidence interval	(4.7,6.5)	(4.4,7.1)	(4.4,6.8)	
N	*4,332*	*2,191*	*2,141*	
Severely acutely malnourished based on WHZ and/or oedema	7.1	7.9	6.3	1.6
95% confidence interval	(6.2,8)	(6.6,9.4)	(5.2,7.6)	
N	*4,294*	*2,174*	*2,120*	

Annex D Survey tools

D.1 Consent forms

CONSENT FORM – HOUSEHOLD HEAD

Hello, my name is [_____] and I work for Oxford Policy Management. You may remember some colleagues of mine that came to your house 3 years ago to collect some basic information.

We are conducting a survey in households across 4 states - Katsina, Jigawa, Kebbi and Zamfara to understand the issue of malnutrition in Northern Nigeria. This research has been authorized by the State Government and Federal Ministry of Health. We also have permission to conduct this research from the local leaders in this community.

Your house was randomly selected for this research and today we would like to ask you questions about your family, your family's health and questions specifically for women and children.

We would like to interview the head of the household and the mothers of children that are under the age of 3 years old. We would also like to measure the height and weight of these women and children. The whole survey should only take about 2.5 hours.

All of the information collected will be confidential.

Do I have your permission to conduct this research?

1 = Yes **2 = No**	**|____|**

Thank you for your time.

INTERVIEWER: If the respondent verbally agrees to the survey, please continue to the next section. If the respondent does not agree, then speak to your supervisor. If there is any issue in the community, please provide the names and contact information of the relevant State Coordinators or Project Director.

OBSERVATIONS

CONSENT FORM – COMMUNITY

Hello, my name is [_____] and I work for Oxford Policy Management. You may remember some colleagues of mine that came to your house 3 years ago to collect some basic information.

We are conducting a survey in households across 4 states - Katsina, Jigawa, Kebbi and Zamfara to understand the issue of malnutrition in Northern Nigeria. This research has been authorized by the State Government and Federal Ministry of Health. We also have permission to conduct this research from the local leaders in this community.

Your community was randomly selected for this research and today we would like to ask you questions about general village characteristics such as the distance to markets, Government and NGO programmes, and any major issues like floods, droughts and insecurity.

The whole interview should only take about 15 minutes.

All of the information collected will be confidential.

Do I have your permission to conduct this research?

1 = Yes **2 = No**	**|____|**

Thank you for your time.

INTERVIEWER: If the respondent verbally agrees to the survey, please continue to the next section. If the respondent does not agree, then speak to your supervisor. If there is any issue in the community, please provide the names and contact information of the relevant State Coordinators or Project Director.

OBSERVATIONS

D.2 Questionnaires

Oxford Policy Management

WINNN Working to Improve Nutrition in Northern Nigeria

ORIE Operational Research and Impact Evaluation

1. *Endline survey* – Household Questionnaire

State ID |__| State name _____

LGA ID |__|__| LGA name _____

Locality ID |__|__|__|__| Locality name _____

EA Number |__|__|__|__| EA name _____

Structure Number |__|__|__| Household Number |__|__|

GPS reading – TO BE FILLED OUT BY SUPERVISOR

ID of GPS unit |__|__|__|__|

GPS coordinates

(please write GPS coordinates in both formats)

	DDD	MM	SS	L
Lon	|__|__|__|	|__|__|	|__|__|	|__|
Lat	|__|__|	|__|__|	|__|__|	|__|

Lon |__|__|__|●|__|__|__|__|__|__|__|

Lat |__|__|__|●|__|__|__|__|__|__|__|

GPS signal strength– no. of bars / precision |__|__| bars |__|__|__| meters

Contact outcomes

Visits	Date	Interviewer	Result	Num. Compl.	Appointment Date	Appointment Time
	DD/MM/YYYY	ID		Qqn	*DD/MM/YYYY*	*HH:MM*
1	|__|__|/|__|__|/2016	|__|__|__|	|__|__|	|__|__|	|__|__|/|__|__|/2016	|__|__|:|__|__|
2	|__|__|/|__|__|/2016	|__|__|__|	|__|__|	|__|__|	|__|__|/|__|__|/2016	|__|__|:|__|__|
3	|__|__|/|__|__|/2016	|__|__|__|	|__|__|	|__|__|	|__|__|/|__|__|/2016	|__|__|:|__|__|
4	|__|__|/|__|__|/2016	|__|__|__|	|__|__|	|__|__|	|__|__|/|__|__|/2016	|__|__|:|__|__|

Result codes:

01 = Completed all questionnaires
02 = Completed some questionnaires
03 = No household member at home or no competent member respondent at home
04 = Entire household absent for extended period of time
05 = Postponed
06 = Refused
07 = Dwelling vacant or address not a dwelling
08 = Dwelling destroyed
09 = Dwelling not found
10 = Other (specify): |_____|

HH questionnaire completed by

Interviewer ID |__|__|__| Interviewer Name _____

Signature _____

Final outcome

Record information after interview completed

		Number of household questionnaires	|__|__|
		Number of child questionnaires completed	|__|__|
HH questionnaire Start time	|__|__| : |__|__|	Number of women questionnaires completed	|__|__|
HH questionnaire End time	|__|__| : |__|__|	Total number of completed questionnaires	|__|__|

Control

Supervisor signature	_____	SC signature	_____	Data entry 1 signature	_____	Data entry 2 signature	_____
Supervisor ID	|__|__|	SC ID	|__|__|	Data entry 1 ID	|__|__|	Data entry 2 ID	|__|__|

Section 0 – Identification of baseline and new household members

	Ask for ALL household members				Movers only			Members not listed at baseline		
	q0_01	q0_02	q0_03	q0_04	q0_05	q0_06	q0_07	0_08	q0_09	q0_10
Line number	**First name**	**Family name**	**Is [NAME] listed at baseline?** 1 = Yes→ **next person** 2 = No→ **q0_08**	**Is [*NAME*] currently a member of this household?** 1 = Yes → **next person** 2 = No → **q0_05** *INTERVIEWER: Please note that the person does not need to be physically present at the time of the survey to be a member of the household since that person could be working, travelling, etc.*	**Why is [NAME] no longer a member of this household?** 1 = Not alive (died) 3 = Moved to another house 4 = Married another family 5 = Temporarily not household member (boarding school, prison, on assignment) → **q0_07** 6 = Has never been a household member 96 = Other. Specify (____) 98 = Don't know	**How long ago did [NAME] die/leave the household or die?** *INTERVIEWER: Write answer in years and months. If answer is less than one year, the number of years should be 0. The number of months must be less than 12 and the number of years can't be greater than 4*	*INTERVIEWER:* If **q_05** is (1) 'Not alive (died)' → **next person. Otherwise, continue.** **Where did [NAME] move to?** 1 = Different house in same community 2 = Different community in same LGA 3 = Different LGA in same state 4 = Outside state 5 = Outside Nigeria 96 = Other. Specify (____) 98 = Don't know → **next person**	**Was [NAME] a member of the household in July 2013?** 1 = Yes → **next person** 2 = No	**Why did [NAME] move into this household?** 1 = New-born → **Next person** 2 = Married in 3 = Relocated in 4 = Was away at baseline 96 = Other. Specify (____) 98 = Don't know	**How long ago did [NAME] join the household?** *INTERVIEWER: Write answer in years and months. If answer is less than one year, the number of years should be 0. Number of months has to be less than 12. A new member who wasn't at baseline couldn't have join before BL. Check!* → **next person**
	HOUSEHOLD DEFINITION: *A person or group of related or un-related persons that live together in the same compound/structure unit who acknowledge the same adult male or female as the head of the household.* *INTERVIEWER: This list will be pre-filled with baseline information. Please add all household members that do not appear in this list*									
01			__|	__|	__|	__|years __|months	__|	__|	__|	__| years __| months
02			__|	__|	__|	__|years __|months	__|	__|	__|	__| years __| months
03			__|	__|	__|	__|years __|months	__|	__|	__|	__| years __| months
04			__|	__|	__|	__|years __|months	__|	__|	__|	__| years __| months
05			__|	__|	__|	__|years __|months	__|	__|	__|	__| years __| months
06			__|	__|	__|	__|years __|months	__|	__|	__|	__| years __| months
07			__|	__|	__|	__|years __|months	__|	__|	__|	__| years __| months
08			__|	__|	__|	__|years __|months	__|	__|	__|	__| years __| months
09			__|	__|	__|	__|years __|months	__|	__|	__|	__| years __| months

10			|__|	|__|	|__|	|__|years |__| months	|__|	|__|	|__|	|__| years |__| months
11			|__|	|__|	|__|	|__|years |__| months	|__|	|__|	|__|	|__| years |__| months
12			|__|	|__|	|__|	|__|years |__| months	|__|	|__|	|__|	|__| years |__| months
13			|__|	|__|	|__|	|__|years |__| months	|__|	|__|	|__|	|__| years |__| months
14			|__|	|__|	|__|	|__|years |__| months	|__|	|__|	|__|	|__| years |__| months

Section A – Household roster

	qa_01	qa_02	ALL HH members		Children 0-6 years – this includes children that are currently 6 years old				
NEW Line number	First name	Family name	qa_03	qa_04	qa_05	qa_06	qa_07	qa_08	qa_09
	INTERVIEWER: *List all household members that were listed at baseline and are still members of the household (q0_03 is 'Yes' and q_0_04 is 'Yes) AND new household members (q0_03 is 'No')*		**Gender** 1 = Male 2 = Female *Consistency check in CAPI if member was at baseline and the gender does not match* between BL and EL: The current gender does not match the gender at baseline. The current gender is: [NAME] Please confirm that the current gender is correct. 1. Current gender correct 2. Current gender NOT correct	**How old is [*NAME*] in completed years?** *INTERVIEWER:* *If respondent **only knows date of birth OR birth year**, use 'age reference table' in* **SHOWCARD #1** 96=96 and more 98 = Don't know *Consistency check in CAPI if member was at baseline:* Are you sure that you RECONFIRM the pre-filled age in QA_04? 1 = Yes / 2 = No	**What is [*NAME*]'s date of birth?** *INTERVIEWER: Write date if respondent knows the birth date.* *If mothers know age of child in months or in a combination of years and months, use* **SHOWCARD #2 or #3**, *otherwise use the **EVENT CALENDAR** to work out date of birth.* *DATE MUST BE FILLED IN!* *98 = don't know day/month* *9998 = don't know year* DD MM YYYY	**Does [*NAME*] have a vaccination card?** 1 = Yes 2 = Yes, but none seen → **qa_09** 3 = No → **qa_09** 98 = Don't know → **qa_09**	*INTERVIEWER: Write date of birth as shown on vaccination card.* *98 = no day/month on the card* *9998 = no year on the card* *If the day or the month of birth is less than 10, write 01, 0* *2, 03 etc* DD MM YYYY	**Was this document obtained within the first 2 months after birth?** 1 = Yes 2 = No 98 = Don't know	*Built-in formula in CAPI: [NAME]'s age in months.* *Use birth date in **qa_07** if document was obtained up to 2 MONTHS after birth – see **qa_08**.* *Otherwise, use birth date in **qa_05** for calculation*
01				__|__|	__|__| __|__| __|__|__|__|	__|__|	__|__| __|__| __|__|__|__|	__|	__|__|
02				__|__|	__|__| __|__| __|__|__|__|	__|__|	__|__| __|__| __|__|__|__|	__|	__|__|
03				__|__|	__|__| __|__| __|__|__|__|	__|__|	__|__| __|__| __|__|__|__|	__|	__|__|
04				__|__|	__|__| __|__| __|__|__|__|	__|__|	__|__| __|__| __|__|__|__|	__|	__|__|
05				__|__|	__|__| __|__| __|__|__|__|	__|__|	__|__| __|__| __|__|__|__|	__|	__|__|
06				__|__|	__|__| __|__| __|__|__|__|	__|__|	__|__| __|__| __|__|__|__|	__|	__|__|
07				__|__|	__|__| __|__| __|__|__|__|	__|__|	__|__| __|__| __|__|__|__|	__|	__|__|
08				__|__|	__|__| __|__| __|__|__|__|	__|__|	__|__| __|__| __|__|__|__|	__|	__|__|
09				__|__|	__|__| __|__| __|__|__|__|	__|__|	__|__| __|__| __|__|__|__|	__|	__|__|
10				__|__|	__|__| __|__| __|__|__|__|	__|__|	__|__| __|__| __|__|__|__|	__|	__|__|
11				__|__|	__|__| __|__| __|__|__|__|	__|__|	__|__| __|__| __|__|__|__|	__|	__|__|
12				__|__|	__|__| __|__| __|__|__|__|	__|__|	__|__| __|__| __|__|__|__|	__|	__|__|
13				__|__|	__|__| __|__| __|__|__|__|	__|__|	__|__| __|__| __|__|__|__|	__|	__|__|
14				__|__|	__|__| __|__| __|__|__|__|	__|__|	__|__| __|__| __|__|__|__|	__|	__|__|
15				__|__|	__|__| __|__| __|__|__|__|	__|__|	__|__| __|__| __|__|__|__|	__|	__|__|

	Ask for ALL household members		Married women only		Women in polygamous marriages only	Ask for ALL household members			
		qa_10	qa_11 A	qa_11	qa_12	qa_13	qa_14	qa_15	qa_16
NEW Line number	*Carry over First name from previous page*	*INTERVIEWER: If [NAME] is less than 10 years old (from qa_13) → **qa_13*** **What is [*NAME*]'s current marital status?** 1 = Married (monogamous) 2 = Married (polygamous) 3 = Divorced / separated → ..... **qa_13** 4 = Never married → ..... **qa_13** 5 = Widowed → ..... **qa_13**	*INTERVIEWER: if answer to question qa_10 is 'Married (monogamous)' or 'Married (polygamous)' and answer to qa_03 is 'female' continue* **Does the husband live in the household?** 1 = Yes 2 = No	**What is the name of [*NAME*]'s husband?** *INTERVIEWER: Write in the NEW line number.* 99 = not in this household	*INTERVIEWER: if answer to question qa_10 is 'Married (polygamous)' continue. Otherwise → **qa_13*** **Is [*NAME*] the first, second, third or fourth wife?** 1 = First wife 2 = Second wife 3 = Third wife 4 = Forth wife	**Does [*NAME*]'s biological mother live in the household?** 1 = Yes 2 = No → **qa_15**	**What is the name of [*NAME*]'s biological mother?** *INTERVIEWER: Write in the new line number.* 99 = not in this household	**Does [*NAME*]'s biological father live in the household?** 1 = Yes 2 =No → **qa_17**	**What is the name of [*NAME*]'s biological father?** *INTERVIEWER: Write in the new line number.* 99 = not in this household
01		__		___	__	__	___	__	___
02		__		___	__	__	___	__	___
03		__		___	__	__	___	__	___
04		__		___	__	__	___	__	___
05		__		___	__	__	___	__	___
06		__		___	__	__	___	__	___
07		__		___	__	__	___	__	___
08		__		___	__	__	___	__	___
09		__		___	__	__	___	__	___
10		__		___	__	__	___	__	___
11		__		___	__	__	___	__	___
12		__		___	__	__	___	__	___
13		__		___	__	__	___	__	___
14		__		___	__	__	___	__	___
15		__		___	__	__	___	__	___

		Members between 4 to 24 years old	ALL household members (built-in checks and skips procedures in CAPI)		
		qa_17	qa_18	qa_19	
NEW Line number	INTERVIEWER: *Carry over First name from previous page*	**Is [*NAME*] currently attending FORMAL school?** 1 = Yes 2 = No	*Summary of women of reproductive age* *Refer to **qa_03** and **qa_04*** *Please cross the box with the new line number of all women between the ages of 15-49. If not woman 15-49 leave blank.* ☒	*Summary of Children under 3* *Refer to **qa_09*** *Please cross the box with the new line number of all children under 3 years old. If not child below 3 leave blank.* ☒	**If there is at least one eligible child (0-3 years old) in the household :** 1) Complete the Household Questionnaire 2) Continue to the Child Interview and make sure you complete one interview per eligible child (0-3 years old) 3) Complete one Woman Interview for every woman who is a mother of at least one eligible child 4) Complete the anthropometric section for each eligible chid and their mother. **If there is at least one eligible child (0-3 years old) in the household AND at least one child 3-6 who was at baseline and is still a member:** 1) Complete the Household Questionnaire 2) Continue to the Child Interview and make sure you complete one interview per eligible child (0-3 years old) 2) Complete one Woman Interview for every woman who is a mother of at least one eligible child 3) Complete the anthropometric section for each eligible chid, her/his mother, and every child 3-6 who was at baseline and it is still a member **If there are 0 eligible child (0-3 years old) in the household BUT at least one child 3-6 who was at baseline and is still a member:** 1) Complete the Household Questionnaire 2) Complete the anthropometric section for all the children who were at baseline and are still members of the household **If there are 0 eligible child (0-3 years old) in the household and 0 children 3-6 who were at baseline and are still a member:** STOP HERE AND CALL YOUR SUPERVISOR
01		⌊_⌋	01 ☐	01 ☐	
02		⌊_⌋	02 ☐	02 ☐	
03		⌊_⌋	03 ☐	03 ☐	
04		⌊_⌋	04 ☐	04 ☐	
05		⌊_⌋	05 ☐	05 ☐	
06		⌊_⌋	06 ☐	06 ☐	
07		⌊_⌋	07 ☐	07 ☐	
08		⌊_⌋	08 ☐	08 ☐	
09		⌊_⌋	09 ☐	09 ☐	
10		⌊_⌋	10 ☐	10 ☐	
11		⌊_⌋	11 ☐	11 ☐	
12		⌊_⌋	12 ☐	12 ☐	
13		⌊_⌋	13 ☐	13 ☐	
14		⌊_⌋	14 ☐	14 ☐	
15		⌊_⌋	15 ☐	15 ☐	

Question number	Question and instructions		Answer
qa_20	**Does the household head live in the household?**	1 = Yes 2 = No → **qa_22**	|__|
qa_21	*INTERVIEWER: Enter the new LINE NUMBER from the household roster in front of his/her name.*		|__|__| new line number → **qa_23**
qa_22	**Who is the person in-charge of the household when household head is not present?** *INTERVIEWER: Enter the new LINE NUMBER from the household roster in front of his/her name.*		|__|__| new line number
qa_23	*INTERVIEWER: Which household member is acting as the main respondent for the household questionnaire?* *Enter his/her new LINE NUMBER from the household roster in front of the name.*		|__|__| new line number

Section B: Household characteristics

Question number	Question and instructions	Answer codes		Answer
EDUCATION AND ECONOMIC ACTIVITY				
qb_01	**What is the household head's occupation, that is, what kind of work does he/she mainly do?** *<u>INTERVIEWER:</u> Do not include shock-responsive strategies, such as selling livestock when cash is needed.*	1 = Farming / herding mainly for subsistence 2 = Commercial farming / herding to sell produce 3 = Paid work	4 = Own business 5 = Not working or inactive ....→**qb_03** 96 = Other (specify) |_____|	|__|__|
qb_02	**Does the household head usually work throughout the year, or does work seasonally, or only once in a while?**	1== Throughout the year 2== Seasonally/ part of the year 3== Once in a while		|__|
qb_03	**What is the highest level of formal education the household head has completed?**	1 = No education 2 = Nursery 3 = Primary Grade 1 4 = Primary Grade 2 5 = Primary Grade 3 6 = Primary Grade 4	7 = Primary Grade 5 8 = Primary Grade 6 9 = Junior secondary 10 = Senior secondary 11 = Tertiary / post-secondary 98 = Don't know	|__|__|
qb_04	**Did the household head attend or is currently attending Islamia education?**	1 = Yes 2 = No 98 = Don' know		|__|__|
qb_05	**Does the head of this household participate in any local groups that people join or attend?** **This could be a voluntary association or religious group that meets outside of worship services.**	1 = Yes 2 = No 98 = Don't know		|__|__|
WATER AND SANITATION				
qb_06	**What is the main source of DRINKING water for members of your household?**	1 = Piped water into dwelling……→ **qb_09** 2 = Piped water to yard/plot ..→ **qb_09** 3 = Public tap/standpipe 4 = Tubewell/borehole 5 = Protected dug well 6 = Unprotected dug well 7 = Protected spring 8 = Unprotected spring 9 = Rainwater collection	10 = Tanker truck..... → **qb_09** 11 = Bottled / sachet..... → **qb_09** 12 = Cart with small tank/drum .....→ **qb_09** 13 = Surface water (river, dam, lake, pond, stream, canal, irrigation channels) 96 = Other (specify) |_____| 98 = Don't know	|__|__|

Question number	Question and instructions	Answer codes	Answer
qb_07	**Where is the source for DRINKING water located?**	1 **=** In own house / dwelling ..... → **qb_09** 2 **=** In own compound / yard ..... → **qb_09** 3 **=** Outside of compound / further away	⌊__⌋
qb_08	**How long does it take to go from your home to the main source of DRINKING water if you were walking?** *<u>INTERVIEWER</u>: This refers to ONE WAY distance and DOES NOT include time waiting to collect the water.*	1 = 0 minutes – less than 30 minutes 2 = 30 minutes – less than 60 minutes 3 = 1 hour – less than 2 hours 4 = 2 hour – less than 5 hours 5 = 5 hours or more 98 = Don't know	⌊__⌋⌊__⌋
qb_09	**Do you treat the water in any way to make it safer to drink?**	1 **=** Yes 2 **=** No ..... → **qb_11** 98 **=** Don't know ..... → **qb_11**	⌊__⌋⌊__⌋
qb_10	**What do you do to treat your drinking water?**	1 **=** Boil 2 **=** Add bleach or chlorine 3 **=** Strain it through a cloth 4 **=** Strain it through a water filter (ceramic, sand, composite) 5 **=** Solar disinfection 6 **=** Let it stand still 96 **=** Other (specify) ⌊_____⌋	⌊__⌋⌊__⌋
qb_11	**Does your household use the same source of water for cooking?**	1 **=** Yes ..... → **0** 2 **=** No	⌊__⌋
qb_12	**What is the main source of water used by your household for cooking?**	1 = Piped water into dwelling 2 = Piped water to yard/plot 3 = Public tap/standpipe 4 = Tubewell/borehole 5 = Protected dug well 6 = Unprotected dug well 7 = Protected spring 8 = Unprotected spring 9 = Rainwater collection 10 = Tanker truck 11 = Bottled / sachet 12 = Cart with small tank/drum 13 = Surface water (river, dam, lake, pond, stream, canal, irrigation channels) 96 = Other (specify) ⌊_____⌋ 98 = Don't know	⌊__⌋⌊__⌋
	CAPI only: if qb_06 OR qb_12 are equal to 3, 4, 5, 6, 7, 8 OR 9 continue. Otherwise, go to qb_14A		
qb_13	**Has this water source been improved in the <u>LAST 3 YEARS</u>?**	1 **=** Yes 2 **=** No 98=Don't know	⌊__⌋⌊__⌋

Question number	Question and instructions	Answer codes	Answer
qb_14A	**What type of toilet facility do members of your household usually use?**	1 = Flush/pour flush 2 = Ventilated improved pit latrine (VIP) 3 = Pit latrine with slab 4 = Pit latrine without slab/open pit 5 = Bucket ..... → **qb_15** 6 = Hanging toilet/hanging latrine → **qb_15** 7 = No facilities or bush or field → **qb_15** 96 **=** Other (specify) |_____| → **qb_15**	|__|__|
	CAPI *only: if qb_14A equals to 1, 2, 3, OR 4 continue. Otherwise, go to qb_15*		
qb_14B	**Has this toilet facility been improved in the <u>LAST 3 YEARS</u>?**	1 **=** Yes 2 **=** No	|__|__|
qb_15	**Have you <u>EVER</u> heard of a WASHCOM?** **These are committees at the village level that help in the maintenance of water points, sanitation facilities and promote good hygiene practices.**	1 **=** Yes 2 **=** No	|__|__|
ASSETS			
qb_16	**How many rooms are there in this household that are used for sleeping** *<u>INTERVIEWER</u>: Count all the rooms used for sleeping for people listed on the household roster.The number of rooms has to be between 1 and 30.*		|__|__| rooms
qb_17	**Does your household have electricity?**	1 = Yes 2 **=** No	|__|

Question number	Question and instructions	Answer codes	Answer
qb_18	**Does any member of this household own any of the following functional assets?**		
	A. Radio	1 = Yes 2 = No	⌊__⌋
	B. Television	1 = Yes 2 = No	⌊__⌋
	C. Satellite television	1 = Yes 2 = No	⌊__⌋
	D. Mobile phone	1 = Yes 2 = No	⌊__⌋
	E. Fridge	1 = Yes 2 = No	⌊__⌋
	F. Mattress / Bed	1 = Yes 2 = No	⌊__⌋
	G. Stove	1 = Yes 2 = No	⌊__⌋
	H. Bicycle	1 = Yes 2 = No	⌊__⌋
	I. Motor-cycle / scooter	1 = Yes 2 = No	⌊__⌋
	J. Animal drawn cart	1 = Yes 2 = No	⌊__⌋
	K. Motor boat	1 = Yes 2 = No	⌊__⌋
	L. Car / truck	1 = Yes 2 = No	⌊__⌋
	M. Electricity Generator	1 = Yes 2 = No	⌊__⌋
	N. Air Conditioner	1 = Yes 2 = No	⌊__⌋
	O. Computer / laptop	1 = Yes 2 = No	⌊__⌋
	P. Electric Iron	1 = Yes 2 = No	⌊__⌋
	Q.	1 = Yes 2 = No	⌊__⌋
	R.	1 = Yes 2 = No	⌊__⌋
qb_19	**Does anybody in this household own any farm land?** *<u>INTERVIEWER</u>: This includes lands for farming, for home consumption or selling and land that is not used.*	1 = Yes 2 = No..... → **qb_21**	⌊__⌋
qb_20	**In the last 12 months did you or anyone from your household harvest this land for own consumption and/or sale?**	1 = Yes 2 = No	⌊__⌋
qb_21	**Do you or anybody in this household own any of the following livestock, herds or farm animals?**	Cows or bulls	1=Yes 2=No ⌊__⌋
		Donkeys, horses or mules	1=Yes 2=No ⌊__⌋
		Camels	1=Yes 2=No ⌊__⌋

Question number	Question and instructions	Answer codes	Answer
qb_22	**What is the main construction material of the <u>floor</u> of the main room of the dwelling?** *<u>INTERVIEWER</u>: If there is more than one material, then ask which material covers the largest area of the floor in the main room.* *The main room is the room of the household head or the guests' room. If none of these are available, ask about the room where the household head slept last night.* *If the respondent is not able to answer, observe.*	1 = Earth / sand / mud 2 = Dung 3 = Wood planks 4 = Parquet / polished wood 5 = Vinyl / asphalt strips 6 = Ceramic tiles 7 = Cement / concrete 96 = Other (specify) |_____| 98 = Don't know	|__|__|
qb_23	**What is the main construction material of the <u>roof</u> of the main room of the dwelling?** *<u>INTERVIEWER</u>: If there is more than one material, then ask which material covers the largest area of the roof in the main room?* *The main room is the room of the household head or the guests' room. If none of these are available, ask about the room where the household head slept last night.* *If the respondent is not able to answer, observe.*	1 = No roof 2 = Mud / mud bricks 3 = Thatch 4 = Sod 5 = Palm/bamboo 6 = Wood planks / beams 7 = Cardboard 8 = Metal / corrugated iron sheets / Zinc 9 = Calamine/cement fibre 10 = Ceramic tiles 11 = Cement 12 = Roofing shingles 96 = Other (specify) |_____| 98 = Don't know	|__|__|
qb_24	**What is the main construction material of the <u>walls</u> of the main room of the dwelling?** *<u>INTERVIEWER</u>: If there is more than one material, then ask which material covers the largest area of the walls in the main room?* *The main room is the room of the household head or the guests' room. If none of these are available, ask about the room where the household head slept last night.* *If the respondent is not able to answer, observe.*	1 = No Walls 2 = Cane/Palm/Trunks 3 = Earth 4 = Bamboo With Mud 5 = Stone With Mud 6 = Adobe (sun baked bricks) 7 = Reused Wood 8 = Wood 9 = Cardboard 10 = Stone / Bricks / Cement blocks / cement 96 = Other (specify) |_____| 98 = Don't know	|__|__|

Question number	Question and instructions	Answer codes	Answer
ACCESS TO HEALTH SERVICES			
qb_25	**How long would it normally take you to walk from your home to the nearest health facility on a one-way trip?** *INTERVIEWER: Refer to the nearest facility where you would find a nurse, midwife, CHEW or doctor, for example PHC health facilities with basic outpatient health services (health centre, health post, clinics) or hospitals.*	1 = 0 minutes – less than 30 minutes 2 = 30 minutes - less than 60 minutes 3 = 1 hour - less than 2 hours 4 = 2 hour - less than 5 hours 5 = 5 hours or more 8 = Don't know	|__|__|
qb_26	**How long would it normally take you to walk from your home to the nearest place where you can buy malaria medicine on a one-way trip?** *INTERVIEWER: Besides PHCs, hospitals and clinics, this question should also refer to chemist, dispensaries, pharmacies, street vendors, etc.*	1 = 0 minutes – less than 30 minutes 2 = 30 minutes - less than 60 minutes 3 = 1 hour - less than 2 hours 4 = 2 hour - less than 5 hours 5 = 5 hours or more 98 = Don't know	|__|__|

Section C: Food Security

Question number	Question text and instructions	Answer options	Answer
qc_01	**During the last 4 weeks, was there ever no food to eat of any kind in your house because of a lack of resources to get food?**	1=Yes 2=No.....→ **qc_03**	|__|
qc_02	**How often did this happen in the last 4 weeks?**	1= Rarely (1-2 times) 2= Sometimes (3-9 times) 3 = Often (more than 10 times)	|__|
qc_03	**During the last 4 weeks, did you or any household member go to sleep hungry because there was not enough food?**	1=Yes 2=No.....→ **qc_05**	|__|
qc_04	**How often did this happen in the last 4 weeks?**	1= Rarely (1-2 times) 2= Sometimes (3-9 times) 3 = Often (more than 10 times)	|__|
qc_05	**During the last 4 weeks, did you or any household member go a whole day and night without eating at all because there was not enough food?**	1=Yes 2=No.....→ **Section D**	|__|
qc_06	**How often did this happen in the last 4 weeks?**	1= Rarely (1-2 times) 2= Sometimes (3-9 times) 3 = Often (more than 10 times)	|__|

SECTION D: Direct Observation

Question number	Question text and instructions	Answer options	Answer
HAND WASHING			
qd_01	**When your hands get dirty, where do you wash your hands?** *INTERVIEWER; If respondent says he/she washes his/her hands anywhere in the compound, for example because he/she washes hands in a basin, answer should be inside the dwelling/compound.*	1 = place inside the dwelling/compound 2 **=** place outside the compound→ **qd_06** 98 = Don't know ..... → **qd_06**	|__|
qd_02	**Can you please show me where you wash your hands?** *INTERVIEWER: If respondent says he/she washes his/her hands anywhere in the compound, for example because he/she washes hands in a basin, ask to show you where they get the water from, even if it is from stored water.*	1 = Yes 2 = No.....→ **qd_06**	|__|
qd_03	*INTERVIEWER OBSERVE: Is there water at the place for hand washing?*	1 = available 2 **=** not available	|__|
qd_04	*INTERVIEWER OBSERVE: Is there soap, detergent, or other cleansing agent (e.g. ash) at the place for hand washing?*	1 = available 2 **=** not available ..... → **qd_06**	|__|
qd_05	*INTERVIEWER OBSERVE: Which type of detergent is present at the place for hand washing?*	1 = Soap or detergent (bar, liquid, powder, paste) 2 = Ash, mud, sand 96 = Other. Specify (_____)	|__|
DRINKING WATER			
qd_06	**Can you please show me where your drinking water is stored?** *INTERVIEWER: Ask respondent to show the place where drinking water is usually stored even though there is no stored water at that moment*	1 = Yes 2 = No..... → **qd_09**	|__|
qd_07	*INTERVIEWER OBSERVE: Is there water at the place where water is usually stored?*	1 = available 2 = not available	|__|

qd_08	*INTERVIEWER OBSERVE: Are drinking water containers covered or not covered?* *If the water containers have covers, but are NOT covered, the answer to this question should be 'Not covered'*	1 = all covered 2 = some covered and some uncovered 3 = none covered	|__|
COOKING			
qd_09	**Can you please show me where food is normally cooked?**	1 = Yes 2 = No..... → **qd_11**	|__|
qd_10	*INTERVIEWER OBSERVE: Is the cooking typically done indoors or outdoors?*	1 = Inside the dwelling 2 = Outside the dwelling, but inside another building/structure 3 = Outdoors	|__|
qd_11	**What type of fuel do you mainly use for cooking?**	1 = electricity 2 = gas 3 = kerosene stove 4 = coal / lignite /charcoal 5 = fire wood 6 = straw / shrubs / grass 7 = animal dung 96 = Other (specify) |_____|	|__|
qd_12	*INTERVIEWER OBSERVE: Are there animals **AROUND** the dwelling/compound?*	1 = Yes 2 = No → **Next questionnaire**	|__|
qd_13	*INTERVIEWER OBSERVE: Are there animals **INSIDE** the dwelling?*	1 = Yes → **Next questionnaire** 2 = No → **Next questionnaire**	|__|

Oxford Policy Management

WINNN Working to Improve Nutrition in Northern Nigeria

ORIE Operational Research and Impact Evaluation

3. |___| *Endline survey* – **Child Questionnaire**

Observations

State ID |__| State name _____

LGA ID |__|__| LGA name _____

Locality ID |__|__|__|__| Locality name _____

EA Number |__|__|__|__| EA name _____

Structure Number |__|__|__| Household Number |__|__|

	NEW line number from HH roster	*Name*
Child	|___|___|	

Questionnaire completed by

Interviewer ID |__|__|__| Interviewer name _____

Interviewer signature _____

Interview date DD |__|__| MM |__|__| **2 0 1 6**

Interview Start time |__|__| : |__|__|

Interview End time |__|__| : |__|__|

This module is to be administered to ALL CHILDREN UNDER THREE YEARS OLD identified in the household roster (refer to QA_24). This module must be answered by the child's mother or main caregiver.
A separate module must be completed for each eligible child.

Control

Supervisor signature	_____	SC signature	_____	Data entry 1 signature	_____	Data entry 2 signature	_____
Supervisor ID	|__|__|	SC ID	|__|__|	Data entry 1 ID	|__|__|	Data entry 2 ID	|__|__|

SECTION E: Infant and Young Child Feeding and Dietary Diversity

Question number	Question and instructions	Answer codes	Answer
qe_01A	**Which household member is acting as the main respondent for the child questionnaire about [CHILD NAME]?**	1 = Mother 2 = Father 3 = Co-wife 4 = Mother in law 5 = Other woman in the household 6 = Other man in the houeshold 96 = Others	☐
qe_01	**I would like to ask you some questions about [CHILD NAME]. Has [CHILD NAME] ever been breastfed?**	1 = Yes .. → Error! Reference source not found. 2 = No 98 = Don't know → Error! Reference source not ound.	|__|
qe_02	**Why was [CHILD NAME] never breastfed?**	1 = Mother/carer was sick .. → Error! Reference ource not found. 2 = Mother/carer was away from the home.. → Error! Reference source not found. 3 = Mother/carer didn't want to breastfeed.... → Error! Reference source not found. 4 = Baby was given other liquids instead, no need to breastfeed..→ Error! Reference source ot found. 96 = Other specify...→ Error! Reference source ot found. 98 = Don't know → Error! Reference source not ound.	|__|
qe_03	**Was [CHILD NAME] breastfed yesterday during the day or at night?**	1 = Yes .. → Error! Reference source not found. 2 = No 98 = Don't know	|__|

Question number	Question and instructions	Answer codes	Answer
qe_04	**Sometimes babies are fed with breast milk in different ways, for example by spoon, cup or bottle. This can happen when the mother cannot always be with her baby. Sometimes babies are breastfed by another woman, or given breast milk from another woman by spoon, cup or bottle or some other way. This can happen if a mother cannot breastfeed her own baby.** **Did [CHILD NAME] consume breast milk in any of these ways yesterday during the day or at night?**	1 = Yes 2 = No 98 = Don't know	|__|
qe_05	*INTERVIEWER: Check pattern of responses in* **Error! Reference source not found.** and **Error! Reference source not found.**. If *[CHILD NAME] HAS EVER been breastfed and WAS NOT breastfed yesterday* (**Error! Reference source not found.**) .....|__1__| (**Error! Reference source not found.**) .....|__2__| ..... (qk_04) |__2, 98__| ask.....→ Error! Reference source not found. If *[CHILD NAME]* (**Error! Reference source not found.**) .....|__1__| (**Error! Reference source not found.**) .....|__2__| ..... (qk_04) |__1__| ask.....→ Error! Reference source not found. If *[CHILD NAME]* (**Error! Reference source not found.**) .....|__2__| (**Error! Reference source not found.**) .....|_____| ..... (qk_04) |__1__| ask.....→ Error! Reference source not found. If *[CHILD NAME] HAS EVER been breastfed and WAS breastfed yesterday* (**Error! Reference source not found.**) .....|__1__| (**Error! Reference source not found.**) .....|__1__| ask .....→ Error! Reference source not found. If *[CHILD NAME] HAS NOT EVER been breastfed* (**Error! Reference source not found.**) .....|__2__| (**Error! Reference source not found.**) .....|_____| ask .....→ Error! eference source not found. *If any answer is Don't know* (**Error! Reference source not found.**) |__98__|..... or (**Error! Reference source not found.**) |__1___| (**Error! Reference source not found.**) .....|__1, 2, 98__|... (**Error! Reference source not found.**) |__98__| ask .....→ Error! Reference source not found.		
qe_06	**Is [CHILD NAME] still being breastfed?**	1 = Yes 2 = No... → Error! Reference source not found.	|__|

Question number	Question and instructions	Answer codes	Answer
qe_07	**Is [CHILD NAME] being fed with <u>other liquids,</u> such as water, or with <u>food</u>?**	1 = Yes ..→Error! Reference source not found. 2 = No .....→Error! Reference source not found.	|__|
qe_08	**For how many months was [CHILD NAME] breastfed for in total?** **This includes both exclusive and non-exclusive breastfeeding.**		|__|__| months
qe_09	**For how long was [CHILD NAME] exclusively breastfed?** *<u>INTERVIEWER</u>: explain exclusively breastfeeding means only giving the child breast milk and nothing else (except medicine is also allowed). If the baby is also given food, water, holy water (zamzam), herbs then it is NOT exclusively breastfed.* *Built-in validations in CAPI:* *This answer can't exceed the age of the child in months.* *This answer can't exceed the number of months for which the child was breastfed IN TOTAL*	00 = If less than 1 month 98 = Don't know	|__|__| months
qe_10	**How long after birth did you first put [CHILD NAME] to the breast?** *<u>INTERVIEWER</u>: Write the answer code number in the space provided.* *Immediately is in the first 1 hour after birth. If respondent says 1 hour, write code '1'* *If more than 1 hour after birth, write code '2' and enter 'hours'.* *If the woman doesn't know the exact number of hours of days, write code '98'.*	1 = 0 – 1 hour (immediately after birth) 2 = More than 1 hour - 24 hours after birth 3 = More than 1 day after birth 98 = Don't know	|__|__|
qe_11	**During the first three days of [CHILD NAME]'s life, was anything fed to [CHILD NAME] other than breast milk, including anything placed inside his/her mouth immediately after the birth?**	1 = Yes 2 = No ....→ Error! Reference source not found. 98 = Don't know ..→ **Error! Reference source ot found.**Error! Reference source not found.	|__|
qe_12	**During the first three days of [NAME]'s life, what was given to [NAME] other than breastmilk?** *<u>INTERVIEWER</u>: <u>**Do not read**</u> the list of answers aloud.* *Please cross the box (☒) for all that apply.* *Keep asking "**anything else?**" until the respondent has mentioned everything she gave to the child in the first days of child's life.* *"Don't Know" can't be combined with other answers.*	01 = Plain water 02 = Infant formula 03 = Milk such as tinned, powdered, or fresh animal milk (do not include breastmilk) 04 = Clear broth 05 = Juice or juice drinks 06 = Yogurt 07 = Thin porridge 08 = Holy water / Islamic water (zamzam) 09 = Honey / dates	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

Question number	Question and instructions	Answer codes	Answer
		10 = Sugar water / glucose water / sugary drinks (e.g. Milo, Ovaltine)	☐
		11 = Traditional herbs / tea / infusions	☐
		12 = Gripe water	☐
		96 = Other (specify) |_____|	☐
		98 = Don't know	☐

Question number	Question and instructions	Answer codes	Answer
	Next I would like to ask you about any <u>liquids</u> you gave or place inside [CHILD NAME]'s mouth	**qe_13** **Did [CHILD NAME] have [ITEM] <u>yesterday</u> during the day or at night?** 1 = Yes 2 = No → **next item** 98 = Don't know → **next item**	**qe_14** **<u>How many times yesterday</u> during the day or at night did [CHILD NAME] consume [ITEM]?**
	01 = Plain water	|__|__|	
	02 = Infant formula	|__|__|	|__|__|
	03 = Milk such as tinned, powdered, or fresh animal milk	|__|__|	|__|__|
	04 = Juice or juice drinks	|__|__|	
	05 = Clear broth	|__|__|	
	06 = Yogurt	|__|__|	|__|__|
	07 = Thin porridge	|__|__|	
	08 = Other liquids (e.g. sugar water, gripe water, tea, etc.)	|__|__|	
qe_15	**Now I would like to ask you about some medicines and vitamins that are sometimes given to infants.** **Was [CHILD NAME] given any vitamin drops or other medicines such as syrups yesterday during the day or at night?**	1 = Yes 2 = No 98 = Don't know	|__|
qe_16	**Was [CHILD NAME] given ORS yesterday during the day or at night?** *<u>INTERVIEWER</u>: Use **[SHOWCARD # 4]** to show ORS sachets.* *Do not include the homemade solution.*	1 = Yes 2 = No 98 = Don't know	|__|

Question number	Question and instructions	Answer codes	Answer
qe_17	**Did [CHILD NAME] EAT any <u>solid, semi-solid, or soft foods YESTERDAY</u> during the day or at night?** **This does not include liquids like breast milk, water.**	1 = Yes 2 = No → Error! Reference source not ound. 98 = Don't know → Error! Reference source ot found.	

qe_18

Now I would like to ask you about EVERYTHING that [CHILD NAME] <u>ATE</u> YESTERDAY during the day or night, whether at home or outside the home.
What did [CHILD NAME] <u>EAT</u> from when he/she woke up in the morning to when he/she went to sleep at night

*<u>INTERVIEWER</u>: When the mother stops probe: "***Did [CHILD NAME] eat anything else?"**
Write down everything the mother says in the first column below.
*<u>After</u> you have made the list, if any dishes are mixed dishes are "***What ingredients were in that (MIXED DISH)?"** *Probe: "***Anything else?***" Continue until respondent says nothing else"*
If the child ate the same food more than once in the last day, please record it once only.

ID	Name of food	Ingredients					Code of ingredients
1		1	2	3	4	5	1|__|__| 2|__|__| 3|__|__| 4|__|__| 5|__|__|
2		1	2	3	4	5	1|__|__| 2|__|__| 3|__|__| 4|__|__| 5|__|__|
3		1	2	3	4	5	1|__|__| 2|__|__| 3|__|__| 4|__|__| 5|__|__|
4		1	2	3	4	5	1|__|__| 2|__|__| 3|__|__| 4|__|__| 5|__|__|
5		1	2	3	4	5	1|__|__| 2|__|__| 3|__|__| 4|__|__| 5|__|__|
6		1	2	3	4	5	1|__|__| 2|__|__| 3|__|__| 4|__|__| 5|__|__|
7		1	2	3	4	5	1|__|__| 2|__|__| 3|__|__| 4|__|__| 5|__|__|
8		1	2	3	4	5	1|__|__| 2|__|__| 3|__|__| 4|__|__| 5|__|__|

Question number	Question and instructions					Answer codes	Answer
	9		1	2	3	4	5	1|__|__| 2|__|__| 3|__|__| 4|__|__| 5|__|__|
	10		1	2	3	4	5	1|__|__| 2|__|__| 3|__|__| 4|__|__| 5|__|__|
	11		1	2	3	4	5	1|__|__| 2|__|__| 3|__|__| 4|__|__| 5|__|__|
	1 = Bread, rice, noodles, porridge, or other foods made from grains [e.g. millet, sorghum, maize, wheat etc.] 2 = Pumpkin, carrots, squash or sweet potatoes that are yellow or orange inside 3 = Irish/White potatoes, white yams, manioc, cassava, or any other foods made from roots 4 = Any dark green, leafy vegetables like spinach, pumpkin leaf etc. 5 = Ripe mangoes, papayas or palm nuts 6 = Any other fruits or vegetables [e.g. bananas, plantains, watermelon, apples, green beans, avocados, tomatoes, onions] 7 = Liver, kidney, heart or other organ meats 8 = Any meat, such as beef, lamb, goat, pork, chicken, or duck 9 = Eggs						10 = Fresh or dried fish or shellfish 11 = Any foods made from any other beans, peas, lentils, or nuts like soy beans, moimoi, akara 12 = Cheese, yogurt or other food made from milk 13 = Red palm oil, foods made with red palm oil, red palm nut, or red palm nut pulp sauce 14 = Any other oil, fats, or butter, or foods made with any of these (e.g. vegetable oil, regular palm oil, NOT red palm oil) 15 = Any sugary foods such as chocolates, sweets, candies, pastries, cakes, or biscuits 18 = Any FORTIFIED BABY FOOD example MADARAR GWANGWANI, CERELAC, SOKOLAC, ASALAC, BUNGULAC 19 = Any other solid, semi-solid, or soft food	
qe_19	**How many times did [CHILD NAME] eat solid, semi-solid, or soft foods YESTERDAY during the day or at night?** **This does not include liquids like breast milk, water.** *INTERVIEWER: Cross-check with dietary recall in Error! Reference source not found.. The nswer to this question should be greater than 0.*							|__|__| times
ANTENATAL CARE								
qe_20	**Now I would like to ask you some questions about when [CHILD NAME]'s mother was pregnant with [CHILD NAME].** **Did she see anybody at a health facility for ANTENATAL CARE during [CHILD NAME]'s pregnancy?**						1 = Yes 2 = No.....→ Error! Reference source not ound. 98 = Don't know .....→ Error! Reference ource not found.	|__|
qe_21	**How many times did [CHILD NAME]'s mother receive ANTENATAL CARE at a health facility during [CHILD NAME]'s pregnancy?** *INTERVIEWER: If the mother said she has received ANC during this pregnancy, the answer to this question should be greater than 0.*							|__| times
qe_22	**When [CHILD NAME]'s mother went to the health facility for ANTENATAL CARE:**							

Question number	Question and instructions	Answer codes	Answer
	A. Did she ever receive any information on breastfeeding, feeding practices and the care of babies [SHOWCARD # 5]?	1 = Yes 2 = No 98 = Don't know	|__|
	B. How many times did she receive this type of counselling? *<u>INTERVIEWER</u>: Probe. Did the mother receive counselling at every antenatal visit for that [CHILD NAME]'S pregnancy? If yes, check that the number of times the mother received training on IYCF practices is equal to the number of times she received ANC.* *If the mother said she has received YICF counselling during ANC, the answer to this quest ion should be greater than 0.*	98 = Don't know	|__|__|
	C. Was [CHILD NAME]'s mother blood pressure measured?	1 = Yes 2 = No 98 = Don't know	|__|
	D. Did [CHILD NAME]'s mother give a urine sample?	1 = Yes 2 = No 98 = Don't know	|__|
	E. Did [CHILD NAME]'s mother give a blood sample?	1 = Yes 2 = No 98 = Don't know	|__|
	F. Was [CHILD NAME]'s mother weight measured?	1 = Yes 2 = No 98 = Don't know	|__|
	G. Was [CHILD NAME]'s mother height measured?	1 = Yes 2 = No 98 = Don't know	|__|
	H. Were [CHILD NAME]'s mother told about things to look out for that might suggest problems with the pregnancy?	1 = Yes 2 = No 98 = Don't know	|__|
qe_23	**<u>During [CHILD NAME]'s pregnancy</u>, did [CHILD NAME]'s mother <u>receive or buy</u> any…?**		
	A. …folic acid / iron folate (iron + folic acid) supplements [SHOWCARD #6]? These are the pills that are given to pregnant women to aid healthy developmental of the baby. It usually comes in yellow colour	1 = Yes 2 = No 98 = Don't know	|__|
	B. …iron supplements [SHOWCARD #7]? These are the pills that increase the quantity of blood and gives strength. It usually comes in red colour	1 = Yes 2 = No 98 = Don't know	|__|
	C. …was she given an injection in the arm to prevent the baby from getting tetanus, that is, convulsions after birth	1 = Yes 2 = No 98 = Don't know	|__|

Question number	Question and instructions	Answer codes	Answer
	D. …any drug for intestinal worms?	1 = Yes 2 = No 98 = Don't know	
	E. …any drugs to keep you from getting malaria?	1 = Yes 2 = No 98 = Don't know	
qe_24	**Did [CHILD NAME]'s mother ALWAYS sleep under a mosquito net while she was pregnant with [CHILD NAME]?**	1 = Yes 2 = No 98 = Don't know	|__|
qe_25	**Was the mosquito net [CHILD NAME]'s mother slept under insecticide treated? [SHOWCARD #8]?** **Insecticide treated mosquito nets are usually obtained for health facilities and should be aired out for at least one day before usage.**	1 = Yes 2 = No 98 = Don't know	|__|
DELIVERY			
qe_26	**Where was [CHILD NAME] born?**	1 **=** At home 2 **=** At a health facility 3 = At the home of a traditional birth attendant (TBA) 96 **=** Other (specify): |_____| 98 = Don't know	|__|
qe_27	**Who assisted during the delivery of [CHILD NAME]?** *INTERVIEWER: Do not read the list of answers aloud.* *Please cross the box (☒) for all that apply.* *Keep asking "**anyone else?**" until the respondent has mentioned everyone that assisted during her delivery*	1 **=** Doctor/Nurse/ Midwife/ Community health extension worker (CHEW) 2 = Traditional birth attendant 3 = Family member 4 = Neighbour 5 = No one assisted 96 = Other. Specify (_____) 98 = Don't know	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

Question number	Question and instructions	Answer codes	Answer
qe_28	**In the two months after [CHILD NAME] was born, did any health care provider or a traditional birth attendant check on his/her health (eg. check cord, baby's temperature, baby feeding well)?** *INTERVIEWER: If the woman gave birth in a health facility, this question should not refer to the care she received at the facility immediately after birth, but to any care she might have received when returning to the facility after having been discharged.*	1 = Yes 2 = No..... → **SECTION F** 98 = Don't know ..... → **SECTION F**	|__|
qe_29	**How long after the birth of [CHILD NAME] did the first check take place?**	1 = 0 - 24 hours (1 day) after birth 2 = More than 1 day after birth or later 98 = Don't know	|__|__|
qe_30	**During these visits, did [CHILD NAME]'s mother receive any information on breastfeeding, feeding practices and the care of babies?**	1 = Yes 2 = No98 = Don't know	|__|

SECTION F: Child Health

Question number	Question and instructions	Answer codes	Answer
qf_01	**Now I would like to ask you some questions about vaccinations that [CHILD NAME] might have had.** **Do you have a Child Health Card where [CHILD NAME]'s vaccinations are written down?** *INTERVIEWER*: sometimes vaccinations are written in a small notebook. This should be considered as a vaccination card.	11 = Yes and Seen 12 = Yes but Not Seen ...→ Error! Reference ource not found. 22 = No ..... → Error! Reference source not ound.	|__|__|
qf_02	*INTERVIEWER: Use the vaccination card to identify which vaccinations the child has had. For each vaccination, if there is a date recorded on the card, then cross the box (☒) at that vaccine, if there is no date recorded then leave the box blank.*		
	A. BCG		☐
	B. OPV 0 - Polio 0 (Polio at birth)		☐
	C. OPV 1 - Polio 1		☐
	D. OPV 2 - Polio 2		☐
	E. OPV 3 - Polio 3		☐
	F. DPT 1		☐
	G. DPT 2		☐
	H. DPT 3		☐
	I. Penta 1		☐
	J. Penta 2		☐
	K. Penta 3		☐
	L. Measles 1 (or MMR)		☐
	M. Measles 2		☐
	N. HBV 0 - Hepatitis B 0 (at birth)		☐
	O. HBV 1 - Hepatitis B 1		☐
	P. HBV 2 - Hepatitis B 2		☐

Question number	Question and instructions	Answer codes	Answer
	Q. HBV 3 - Hepatitis B 3		☐
	R. PCV 1 - pneumococcal conjugate		☐
	S. PCV 2 - pneumococcal conjugate		☐
	T. PCV 3 - pneumococcal conjugate		☐
	U. Yellow Fever		☐
	V. Vitamin A 1st Dose		☐
	W. Vitamin A 2nd Dose.....		☐
	X. Conjugate A CSM		☐
	Y. Rota 1		
	Z. Rota 2		
	AA. IPV		
qf_03	**Has [CHILD NAME] had any vaccinations that are not recorded on this card, including vaccinations given in a national immunization day campaign?**	1 = Yes .....→ Error! Reference source not ound. 2 = No.....→ Error! Reference source not ound. 98 = Don't Know → Error! Reference source ot found.	⌊_⌋
qf_04	**Did [CHILD NAME] ever have any vaccinations to prevent him/her from getting diseases, including vaccinations received in a national immunization day campaign**	1 = Yes 2 = No.....→ Error! Reference source not ound.	⌊_⌋

Question number	Question and instructions	Answer codes	Answer
qf_05	**Please tell me if [CHILD NAME] had any of the following vaccinations.** **(Apart from the BCG vaccinations recorded on the card) Has [CHILD NAME] ever received a BCG vaccination (other BCG vaccinations) against tuberculosis, that is, an injection in the arm or shoulder that usually causes a scar?** *INTERVIEWER: If any BCG vaccination was recoded (with date) on the vaccination card, this question should refer to any BCG vaccination **on top** of the one(s) recorded on the vaccination card.*	1 = Yes 2 = No 98 = Don't Know	⌊__⌋
qf_06	**(Apart from the polio vaccines recorded on the card) Has [CHILD NAME] ever received a Polio vaccine (other polio vaccines), that is, drops in the mouth?**	1 = Yes 2 = No.....→ Error! Reference source not ound. 98 = Don't Know → Error! Reference source ot found.	⌊__⌋
qf_07	*(**For CAPI only**: if ql_02_B = OPV 0 - Polio 0 is marked, then skip this question)* **Was the first polio vaccine given in the first two weeks after birth or later?**	1 = Within first two weeks after birth 2 = More than two weeks after birth 98 = Don't Know	⌊__⌋
qf_08	**How many times was the polio vaccine given in total?** *INTERVIEWER: Write '96' if answer is 'too many to count '. Write '98' if answer is 'don't know '. If respondent said the child has received polio vaccine, the answer to this question should be greater than 0.*	96 = too many to count 98 = Don't Know	⌊__⌋__⌋ times
qf_09	**(Apart from the DPT or Penta vaccinations recorded on the card) Has [CHILD NAME] ever received a DPT or Penta vaccination (other DPT or Penta vaccinations) - that is, an injection given in the thigh or buttocks to prevent him/her from getting tetanus, whooping cough or diphtheria?**	1 = Yes 2 = No.....→ Error! Reference source not ound. 98 = Don't Know → Error! Reference source ot found.	⌊__⌋

Question number	Question and instructions	Answer codes	Answer
qf_10	**How many times was the DPT or Penta vaccination given?** *INTERVIEWER: Write '98' if answer is 'don't know '. If respondent said the child has received DPT vaccination, the answer to this quest ion should be greater than 0. Out of range: double-check the respondent is not including other vaccinations in the total count.*	98 = Don't Know	|__| times
qf_11	**(Apart from the Hepatitis B vaccination recorded on the card) Has [CHILD NAME] ever received a Hepatitis B vaccination (other Hepatitis B vaccinations) - that is, an injection in the thigh or buttocks to prevent him/her from getting Hepatitis B?** *INTERVIEWER: Probe by indicating that the Hepatitis B vaccine is sometimes given at the same time as the polio and DPT vaccines.*	1 = Yes 2 = No.....→ Error! Reference source not ound. 98 = Don't Know → Error! Reference source ot found.	|__|
qf_12	**Was the first Hepatitis B vaccine received within the first 24 hours after birth?**	1 = Yes 2 = No 98 = Don't Know	|__|
qf_13	**How many times was a Hepatitis B vaccine received?** *Write '98' if answer is 'don't know '. If respondent said the child has received Hepatitis B vaccination, the answer to this quest ion should be greater than 0. Out of range: double-check the respondent is not including other vaccinations in the total count.*	98 = Don't Know	|__| times
qf_14	**(Apart from the measles or MMR injections recorded on the card) Has [CHILD NAME] ever received a measles injection or an MMR injection (other measles or MMR injections) - that is, a shot in the arm at the age of 9 months or older - to prevent him/her from getting measles?**	1 = Yes 2 = No 98 = Don't Know	|__|
qf_15	**(Apart from the yellow fever vaccination recorded on the card) Has [CHILD NAME] ever received the yellow fever vaccination (other yellow fever vaccinations) – that is, a shot in the arm at the age of 9 months or older – to prevent him/her from getting yellow fever?** *INTERVIEWER: Probe by indicating that the yellow fever vaccine is sometimes given at the same time as the measles vaccine*	1 = Yes 2 = No 98 = Don't Know	|__|

Question number	Question and instructions		Answer codes		Answer
SUPPLEMENTS AND MEDICINE					
	I would like to ask you questions about supplements / medicines given to [CHILD NAME] at home or outside the home.				
		Supplement / medicine / supply	**qf_16 Was [CHILD NAME] <u>ever</u> given…?** 1 = Yes 2 = No → **next item** 98 = Don't Know →**next item**	**qf_17 Was [CHILD NAME] given in the <u>last 6 months</u>…?** 1 = Yes 2 = No 98 = Don't Know	**qf_18 Where did [CHILD NAME] receive this…the <u>last time</u>?** 1 = At own house 2 = At the community 3 = At the health facility (includes MNCH weeks and CMAM day) 3 = Chemist 4 = Bought from a street hawker 96 = Other (specify) |_____| 98 = Don't Know
	1	Vitamin A drops [**SHOWCARD # 9**]	|__|	|__|	|__|
	2	Deworming medication. These are drugs for intestinal worms [**SHOWCARD # 10**]	|__|	|__|	|__|
	3	Oral Rehydration Salts (ORS) sachets [**SHOWCARD # 4**]	|__|	|__|	|__|
	4	RUTF/ MADARAR KWAMAZO[**SHOWCARD # 11**]	|__|	|__|	|__|
GENERAL HEALTH					
qf_19	**Has the weight, length and/or height of [CHILD NAME] <u>EVER</u> been measured by a health professional?** *<u>INTERVIEWER</u>: A health professional includes a doctor, nurse, midwife, CHEW or CV*		1 = Yes 2 = No → Error! Reference source not found. 98 = Don't Know ..→ Error! Reference source ot found.		|__|
qf_20	**Has the weight, length and/or height of [CHILD NAME] been measured by a health professional <u>IN THE LAST 6 MONTHS</u>?** *<u>INTERVIEWER</u>: A health professional includes a doctor, nurse, midwife, CHEW or CV.*		1 = Yes 2 = No 98 = Don't Know		|__|

Question number	Question and instructions	Answer codes	Answer
qf_21	**Do you own any insecticide treated mosquito nets?** **Insecticide treated mosquito nets are usually obtained for health facilities and should be aired out for at least one day before usage.**	1 = Yes 2 = No 98 = Don't Know	|__|
qf_22	**Did [CHILD NAME] sleep under <u>ANY</u> mosquito net last night?** *<u>INTERVIEWER:</u> This question refers to **ANY** mosquito net, not only insecticide treated nets.*	1 = Yes 2 = No 98 = Don't Know	|__|
qf_23	**Has [CHILD NAME] had diarrhoea in the <u>LAST 2 WEEKS</u>?**	1 = Yes 2 = No → Error! Reference source not found.	|__|
qf_24	**At any time during the diarrhoea, was there any blood in the stools?**	1 = Yes 2 = No 98 = Don't know	|__|
qf_25	**Did you seek advice or treatment for the diarrhoea from any source in the <u>LAST 2 WEEKS</u>?**	1 = Yes 2 = No → Error! Reference source not found.	|__|
qf_26	**Where did you seek advice or treatment?** *<u>INTERVIEWER</u>:* *Do not read the list of answers aloud.* *Please cross the box (☒) for all that apply.* *Keep asking "**everywhere else?**" until the respondent has mentioned all places she consulted for treatment*	1 = Neighbour, family member, friend	☐
		2 = Traditional practitioner	☐
		3 = Dispensary / Chemist / Shop	☐
		4 = Private medical clinic	☐
		5 = Primary health centre / Health post / mobile clinic	☐
		6 = Hospital	☐
		96 = Other. Specify (_____)	☐
	<u>For CAPI only:</u> *If Error! Reference source not found. = 4, 5 AND/OR 6 continue, otherwise skip to → Error! Reference source not found.*		
qf_27	**When [CHILD NAME] had diarrhoea in the LAST 2 WEEKS, how many times have you taken him/her to the clinic, health centre and/or hospital to seek treatment?** *INTERVIEWER: The answer to this question should be greater than 0. Write '98' if answer is 'don't know'*	98 = Don't know	|__|__| times

Question number	Question and instructions	Answer codes	Answer
qf_28	**At any time during the diarrhoea, was [CHILD NAME] given:** **a) A fluid made from a special packet called ORS?** **b) A government-recommended homemade fluid – SSS (Salt Sugar Solution)?**	1 = Yes 2 = No 98 = Don't know	a) |__| b) |__|
qf_29	**Has [CHILD NAME] been ill with fever at any time in the LAST 2 WEEKS?**	1 = Yes 2 = No → Error! Reference source not ound. 98 = Don't know → Error! Reference ource not found.	|__|
qf_30	**At any time during the illness, did [CHILD NAME] have blood taken from his/her finger or heel for testing?**	1 = Yes 2 = No 98 = Don't know	|__|
qf_31	**Has [CHILD NAME] had an illness with cough at any time in the LAST 2 WEEKS?**	1 = Yes 2 = No 98 = Don't know	|__|
	INTERVIEWER: Check pattern of responses in ***Error! Reference source not found.*** *and* ***Error! Reference source not found.*** *If* ***Error! Reference source not found.*** *is equal to 'Yes', continue with question* ***Error! Reference source not found..*** *If* ***Error! Reference source not found.*** *is equal to 'No' or 'Don't know' and* ***Error! Reference source not found.*** *is equal to 'Yes'* → Error! Reference source not nd. *If both* ***Error! Reference source not found.*** *AND* ***Error! Reference source not found.*** *are equal to 'No' or 'Don't know'* → Error! Reference source not found.		
qf_32	**When [CHILD NAME] had an illness with a cough, did he/she breathe faster than usual with short, rapid breaths or have difficulty breathing?**	1 = Yes 2 = No → Error! Reference source not ound. 98 = Don't know → Error! Reference source ot found.	|__|
qf_33	**Was the fast or difficult breathing due to a problem in the chest or to a blocked or runny nose?**	1 = Chest only 2 = Nose only 3 = Both 96 = Other (specify) 98 = Don't know	|__|
qf_34	**Did you seek advice or treatment for the illness from any source in the LAST 2 WEEKS?**	1 = Yes 2 = No → Error! Reference source not found.	|__|
qf_35		1 = Neighbour, family member, friend	☐

Question number	Question and instructions	Answer codes	Answer
	Where did you seek advice or treatment? *INTERVIEWER: Do not read the list of answers aloud.* *Please cross the box (☒) for all that apply.* *Keep asking "**everywhere else?**" until the respondent has mentioned all places she consulted for treatment*	2 = Traditional practitioner	☐
		3 = Dispensary / Chemist / Shop	☐
		4 = Private medical clinic	☐
		5 = Primary health centre / Health post / mobile clinic	☐
		6 = Hospital	☐
		96 = Other. Specify (_____)	☐
	For CAPI only: *If Error! Reference source not found. = 4, 5 AND/OR 6 continue, otherwise skip to → Error! Reference source not found.*		
qf_36	**How many times have you taken [CHILD NAME] to the clinic, health centre and/or hospital to seek treatment for this illness in the LAST 2 WEEKS?***INTERVIEWER: The answer to this quest ion should be greater than 0*	98 = Don't know	|__|__| times
qf_37	**In the last 6 months, has [CHILD NAME] been hospitalised for an overnight stay in a hospital?** *INTERVIEWER: If the child is less than 6 months, refer this question to the period since the child was born. Hospitalization for delivery should not be considered in this question.*	1 = Yes 2 = No → **Next section**	|__|
qf_38	**How many times in the last 6 months has [LAST CHILD] been hospitalised?** *INTERVIEWER: Answer should be greater than 0*	98 = Don't know	|__| times
qf_39	**The last time [CHILD NAME] was hospitalised, how many nights in total did he/she spend in the hospital?** *INTERVIEWER: Answer should be greater than 0*	98 = Don't know	|__| days
qf_40	**The last time [CHILD NAME] was hospitalised, how much money did your household spend in total for medicines, fees, travel, accommodation, etc.?** *INTERVIEWER: Ask the respondent to estimate the total expenses. 00 = Did not spend any money 97 = Refused to answer 98 = Don't know*	000000 = nothing 999997 = Refusal 999998 = Don't Know.....Naira	|__|__|__|__|__|__|

SECTION G: Community-based Management of Acute Malnutrition

Question number	Question and instructions	Answer codes	Answer
qg_01	**Has somebody EVER measured the mid-upper arm circumference (MUAC) of [CHILD NAME using medical equipment like this. [SHOWCARD #12]**	1 = Yes 2 = No.....→ Error! Reference source not ound.	|__|
qg_02	**How many times in the LAST 6 MONTHS was this done?** *INTERVIEWER: Probe. **[Use CHILD Show Card #2]** to show what the MUAC measuring tape looks like. Write '98' if answer is 'don't know '.*	98 = Do not know	|__|__| times
	For CAPI only: *If Error! Reference source not found. > 00 continue, otherwise skip to → Error! Reference source not found.*		
qg_03	**How many times in the LAST 30 DAYS?** *INTERVIEWER: The number of times this examination was done in the last 30 days should not exceed the number of times the examination was done in the last 6 months. Write '98' if answer is 'don't know '.*	98 = Do not know	|__| times
qg_04	**THE LAST TIME [CHILD NAME] received this examination, where did it take place?**	1 = At your house 2 = At the community (e.g. at community leader's house, primary school, ceremonies, etc.) 3 = At the health facility 96 = Other Specify (_____) 98 = Don' know	|__|
qg_05	**THE LAST TIME [CHILD NAME] received this examination, were you told to go to a health facility for treatment?**	1 = Yes 2 = No 98 = Don't know	|__|

qg_06	**Has [CHILD NAME] EVER received treatment with MADARAR KWAMAZO (RUTF), Therapeutic Milk (F75/F100) or KWASH PAP at a health facility?** *INTERVIEWER:* Use **SHOWCARD # 11** and **SHOWCARD # 13**. KWASH PAP is known as a local recipe to treat malnutrition. Include outpatient or inpatient (hospitalisation) services	1 = Yes 2 = No→ **Next Questionnaire**	|__|
OUTPATIENT THERAPEUTIC TREATMENT			

qg_07	**Have you <u>EVER</u> taken [CHILD NAME] to a health facility to receive RUTF or KWASH PAP and <u>did NOT stay in a hospital bed overnight</u>?** *<u>INTERVIEWER</u>: We are looking for children that have received treatment in outpatient facilities, NOT HOSPITALISED*	1 = Yes 2 = No..→ Error! Reference source not found.	|__|
qg_08	**When you took your child to the health facility and <u>did not stay overnight in a hospital bed</u>, was the child <u>EVER</u> given RUTF?** *<u>INTERVIEWER:</u>* Use **SHOWCARD # 11**	1 = Yes → Error! Reference source not ound. 2 = No 98 = Don't know	|__|
qg_09	**<u>THE LAST TIME</u> your child was treated with KWASH pap, how many times did you go to the health facility to receive this treatment?** *<u>INTERVIEWER</u>: We are looking for children that have received treatment in outpatient facilities, NOT HOSPITALISED. Write '98' if answer is 'don't know '*	98 = Don't know	|__|__| times
qg_10	**If you had to walk to this health facility where [CHILD NAME] received this treatment, how long would it take you to walk a one-way journey from your home?** *<u>INTERVIEWER</u>: Even if the respondent did not walk to the facility, please ask her to estimate the time it would take to walk. You can probe using the answer codes.*	1 = 0 minutes – less than 30 minutes 2 = 30 minutes – less than 60 minutes 3 = 1 hour – less than 2 hours 4 = more than 2 hours 98 = Don't know	|__|
qg_11	**<u>THE LAST TIME</u> you made this trip, how much money did you spend to travel to this health facility for a one-way journey?** <u>INTERVIEWER</u>: *Write amount in Naira. Write '000000' if they did not spend anything* Write amount in Naira: Write 0 for Nothing Type 97 if Refused Type 98 if Don't Know	0 = nothing 97 = Refusal 98 = Don't Know Naira	|__|__|__|__|__|__| → Error! Reference ource not found.
qg_12	**<u>THE LAST TIME</u> your child was treated with RUTF, how many times did you go to the health facility to receive this treatment?** *<u>INTERVIEWER</u>: We are looking for children that have received treatment in outpatient facilities, NOT HOSPITALISED. The answer to this question should be greater than 0. Write '98' if answer is 'don't know '*	98 = Don't know	|__|__| times
qg_13	**How long ago was <u>THE LAST TIME</u> you went to the health facility to get RUTF and <u>did NOT stay overnight at the facility</u>?** *INTERVIEWER: Write answer in YEARS and MONTHS. Write '98' if answer is 'Don't know' '0' if 'Child is currently hospitalised for malnutrition'. The number of months should not exceed 12.*	98 = Don't know	|__|__| years |__|__| months

qg_14	**THE LAST TIME [CHILD NAME] got treatment with RUTF, did you receive any information on breastfeeding, feeding practices and the care of babies at the health facility?**	1 = Yes, group counselling 2 = Yes, individual counselling 3= Yes, both group and individual counselling 4 = No 98 = Don't know	|__|
qg_15	**If you had to walk to the health facility where [CHILD NAME] received this treatment with RUTF, how long would it take you to walk a one-way journey from your home?** *INTERVIEWER: Even if the respondent did not walk to the facility, please ask her to estimate the time it would take to walk. You can probe using the answer codes.*	1 = 0 minutes – less than 30 minutes 2 = 30 minutes – less than 60 minutes 3 = 1 hour – less than 2 hours 4 = more than 2 hours 98 = Don't know	|__|
qg_16	**The last time you made this trip, how much money did you spend to travel to this health facility for a one-way journey?** *INTERVIEWER: Write amount in Naira. Write '000000' if they did not spend anything* *Write amount in Naira: Write '0' for 'Nothing', '97' for 'Refusal' and '98' if answer is 'Don't know'*	0 = nothing 97 = Refusal 98 = Don't Know	|_____| Naira
qg_17	**Does the RUTF need any preparation, like cooking or mixing it with water, before it can be fed to [CHILD NAME]?**	1 = Yes, needs preparation 2 = No, it does not need preparation 98 = Don't know	|__|
qg_17A	**The last time you gave RUTF to [CHILD NAME] did you prepare it in any way, for example by cooking it or mixing it with water?**	1 = Yes, cooked it / or mixed it with water or other liquids 2 = No, gave directly with no preparation 98 = Don't know	
qg_18	**Do you share the MADARAR KWAMAZO (RUTF) sachets you get from the health facility with your other children or anyone else?**	1 = Yes 2 = No	|__|
qg_19	**Is it possible to buy MADARAR KWAMAZO (RUTF) outside of the health facility?**	1 = Yes 2 = No → Error! Reference source not found.	|__|
qg_20	**From whom can people buy MADARAR KWAMAZO (RUTF)?** *INTERVIEWER:* *Do not read the list of answers aloud.* *Please cross the box (☒) for all that apply.*	1 = Market sellers / street hawkers / shops 2 = Caregivers of child prescribed with RUTF 3 = Health workers 4 = Community volunteers 96 = Others. Specify (_____) 98 = Do not know	☐ ☐ ☐ ☐ ☐ ☐
INPATIENT THERAPEUTIC TREATMENT (hospitalisation services)			
qg_21	**For this special treatment, did [CHILD NAME] ever stayed in a hospital bed overnight at a health facility?**	1 = Yes 2 = No→ **Next questionnaire**	|__|

qg_22	**<u>In the last 6 months</u>, how many times has [CHILD NAME] been hospitalised to receive this special treatment for malnutrition?** *INTERVIEWER: Write '98' if answer is 'Don't know'*	98 = Don't know	|__| times
qg_23	**<u>How long ago was the last time</u> [CHILD NAME] got hospitalised to receive this special treatment for malnutrition?** *INTERVIEWER: Write answer in YEARS and MONTHS. If answer is less than one year, the number of years should be '00'. If answer is less than 10 years or less the n 10 months, type '0' followed by the number of years. Write '98' if answer is 'Don't know'. The number of months should no exceed 12. This answer is not coherent w ith the number of hospitalisation in the last 6 months*	00 = if child is current hospitalised for malnutrition 98 = Don't know	|__|__| years |__|__| months
	<u>For CAPI only:</u> If **qg_22>0 AND qg_23> 6 months, then display ERROR message**		
qg_24	**THE LAST TIME [CHILD NAME] was hospitalised, <u>who</u> told you to bring your child to the health facility?**	01 = Nobody / Went by herself 02 = A health worker 03 = A community volunteer 96 = Other Specify (_____) 98 = Don't know	|__|
qg_25	**THE LAST TIME [CHILD NAME] was hospitalised, <u>how many nights in total</u> did he/she spend overnight in the hospital?** *<u>INTERVIEWER</u>: If child is currently in the hospital, ask for how many days child has been hospitalised so far. Write '98' if answer is 'Don't know'. The answer to this question should be greater than 0*	98 = Don't know	|__| days
qg_26	**THE LAST TIME [CHILD NAME] was hospitalised, <u>what was given</u> to treat malnutrition?** *<u>INTERVIEWER:</u>* Use **SHOWCARD # 13** and **SHOWCARD # 11**. *<u>If Don't Know is selected, then no others opt ion can be picked.</u>*	01 = RUTF / Plumpy'nuts (MADARAR KWAMAZO)	☐
		02 = Therapeutic milk – F75/F100	☐
		03 = Kwash pap	☐
		04 = Other therapeutic food (Specify)|_____|	☐
		96 = Other. Specify |_____|	☐
		98 = Do not know	☐

Oxford Policy Management

WINNN Working to Improve Nutrition in Northern Nigeria

ORIE Operational Research and Impact Evaluation

2. |___| *Endline survey* – **Woman Questionnaire**

Observations

State ID	|__|	State name	_____
LGA ID	|__|__|	LGA name	_____
Locality ID	|__|__|__|__|	Locality name	_____
EA Number	|__|__|__|__|	EA name	_____
Structure Number	|__|__|__|	Household Number	|__|__|

New line number from HH roster *Name*

Woman |___|___|

Questionnaire completed by

Interviewer ID |__|__|__| Interviewer name _____

Interviewer signature _____

Interview date	DD |__|__| MM |__|__| **2 0 1 6**	Interview Start time	|__|__| : |__|__|
		Interview End time	|__|__| : |__|__|

This module is to be administered to ALL MOTHERS OF CHILDREN 0-3 YEARS OLD identified in the household roster. A separate module must be completed for each eligible woman.

Control

Supervisor signature	_____	SC signature	_____	Data entry 1 signature	_____	Data entry 2 signature	_____
Supervisor ID	|__|__|	SC ID	|__|__|	Data entry 1 ID	|__|__|	Data entry 2 ID	|__|__|

SECTION H: Maternal health

Question number	Question and instructions	Answer codes	Answer
	WOMAN CHARACTERISTICS		
qh_01	*INTERVIEWER:* *Is this woman married?* *Refer to Section A of Household questionnaire for Marital Status*	1 = Married (monogamous) 2 = Married (polygamous) 3 = Common law ..... → Error! Reference source not found. 4 = Divorced / separated 5 = Never married..... → Error! Reference source not found. 6 = Widowed	|__|
qh_02	**At what age did you first marry?** *INTERVIEWER: If woman has been married for several times, then ask for first marriage.* *Write age in completed years.* *Age at first marriage should be lower than or equal to respondent's current age.*	98 = Don't know	|__|__|
qh_03	**At what age did you first give birth to a live child – even if he or she lived only a few minutes or hours?** *INTERVIEWER: If woman does not know at what age she gave birth for the first time, use the age of the first born and/or the time between marriage and the first birth to estimate the age. Age at first birth should be lower than or equal to respondent's current age.*		|__|__|
qh_04	**How many live children have you ever given birth to – even if they lived only a few minutes or hours after birth?** *INTERVIEWER: The total number of children should be greater or equal than the number of children that were indicated of the children of this woman in the household roster.*		|__|__| children
qh_04A	**Are you currently pregnant?**	1 = Yes 2 = No 98 = Don't know	|__|__|
qh_05	*INTERVIEWER: Is this woman the household head/ or person in-charge of the household?* *Refer to NEW LINE NUMBER in cover page and HH grid **qa_26** or **qa_27***	1 = Yes ..... → Error! Reference source not found. 2 = No	|__|

Question number	Question and instructions	Answer codes	Answer
qh_06	**Now I would like to ask you questions about your daily life.** **What is your occupation, that is, what kind of work do you mainly do?**	1 = Farming / herding mainly for subsistence 2 = Commercial farming / herding to sell produce 3 = Paid work 4 = Own business 5 = Not working or inactive.....→ Error! eference source not found. 96 = Other (specify) |_____|	|__|
qh_07	**Do you usually work throughout the year, or do you work seasonally, or only once in a while?**	1== Throughout the year 2== Seasonally/ part of the year 3== Once in a while	|__|
qh_08	**What is the highest level of formal education that you completed?**	01 = No education 02 = Nursery 03 = Primary Grade 1 04 = Primary Grade 2 05 = Primary Grade 3 06 = Primary Grade 4 07 = Primary Grade 5 08 = Primary Grade 6 09 = Junior secondary 10 = Senior secondary 11 = Tertiary / post-secondary	|__|__|
qh_09	**Did you attend or are currently attending Islamia education?**	1 = Yes 2 = No 98 = Don't know	|__|
qh_10	**Do you participate in any local groups or associations?** **This could be a voluntary association such as Adashe (women's savings and credit groups) or religious group that meets outside of worship services**	1 = Yes 2 = No 98 = Don't know	|__|
qh_11	*INTERVIEWER: For the following set of questions (including **Section F**) try to move somewhere private.* ***AFTER** you tried to move somewhere private. Please **observe and record** all people who are **STILL** present while you are asking these questions* *Cross the box (☒) for all that apply.*	1 = man/men	☐
		2 = woman/women	☐
		3 = older children (approximately more than 5 years)	☐
		4 = nobody	☐
		*(**note for CAPI:** remove 'Other' category)*	

SECTION I: Intra-Household Decision Making

Question number	Question text and instructions	Answer options	Answer
qi_01A	*INTERVIEWER:* *For the following set of questions (including **Section F**) try to move somewhere private.* ***AFTER** you tried to move somewhere private. Please **observe and record** all people who are **STILL** present while you are asking these questions* *Cross the box (☒) for all that apply.* *If NOBODY is selected, then no other option can be picked*	1 = man/men	☐
		2 = woman/women	☐
		3 = older children (approximately more than 5 years)	☐
		4 = nobody	☐
		*(**note for CAPI:** remove 'Other' category)*	
qi_01	**Do you do any work to earn CASH (for example grow food to sell, sell other things, run a small business, work on a farm or work for someone else)?** *INTERVIEWER: Do not include in-kind payments. Only answer if the answers to qe_06 are 1,2,3,4 or 96.*	1 = Yes 2 = No ..... → Error! Reference source not found.	|____|
qi_02	**Do you usually decide how to spend the money you earn alone, jointly with someone else, or does someone else decide for you?**	1 = Decides alone →**qi_03** 2 = Decides jointly or in consultation with someone else →**qi_02A** 3 = Other person decides. →**qi_02B**	|____|
qi_02a	**Who do you usually make these decisions with?** *INTERVIEWER: move to qf_03*	01 = Co-wife	☐
		02 = Husband	☐
		03 = Mother-in-law	☐
		04 = Father-in-law	☐
		05 = Own mother	☐
		06 = Own father	☐
		07 = Other woman in the household	☐

		08 = Other men in the household	☐
		96 = Other	☐
qi_02b	**Who usually decides for you?** *INTERVIEWER: move to qf_03*	01 = Co-wife	☐
		02 = Husband	☐
		03 = Mother-in-law	☐
		04 = Father-in-law	☐
		05 = Own mother	☐
		06 = Own father	☐
		07 = Other woman in the household	☐
		08 = Other men in the household	☐
		96 = Other	☐
qi_03	**Do you usually make decisions about the health care of your child/children alone, jointly with someone else, or does someone else decide for you?**	1 = Decides alone →**qi_04** 2 = Decides jointly or in consultation with someone else →**qi_03A** 3 = Other person decides →**qi_03B**	|____|
qi_03a	**Who do you usually make these decisions with?** *INTERVIEWER: move to qf_04*	01 = Co-wife	☐
		02 = Husband	☐
		03 = Mother-in-law	☐
		04 = Father-in-law	☐
		05 = Own mother	☐
		06 = Own father	☐
		07 = Other woman in the household	☐
		08 = Other men in the household	☐
		96 = Other	☐
qi_03b	**Who usually decides for you?** *INTERVIEWER: move to qf_04*	01 = Co-wife	☐
		02 = Husband	☐
		03 = Mother-in-law	☐
		04 = Father-in-law	☐
		05 = Own mother	☐
		06 = Own father	☐
		07 = Other woman in the household	☐
		08 = Other men in the household	☐

		96 = Other	☐
qi_04	**Do you usually make decisions about major household purchases alone, jointly with someone else, or does someone else decide for you?** *INTERVIEWER:* These purchases are non-food related **such as mattress, furniture, etc.**	1 = Decides alone →**qi_05** 2 = Decides jointly or in consultation with someone else. →**qi_04A** 3 = Other person decides →**qi_04B**	|____|
qi_04a	**Who do you usually make these decisions with?** *INTERVIEWER: move to qf_05*	01 = Co-wife	☐
		02 = Husband	☐
		03 = Mother-in-law	☐
		04 = Father-in-law	☐
		05 = Own mother	☐
		06 = Own father	☐
		07 = Other woman in the household	☐
		08 = Other men in the household	☐
		96 = Other	☐
qi_04b	**Who usually decides for you?** *INTERVIEWER: move to qf_05*	01 = Co-wife	☐
		02 = Husband	☐
		03 = Mother-in-law	☐
		04 = Father-in-law	☐
		05 = Own mother	☐
		06 = Own father	☐
		07 = Other woman in the household	☐
		08 = Other men in the household	☐
		96 = Other	☐
qi_05	**Do you usually make decisions about what food to buy alone, jointly with someone else, or does someone else decide for you?**	1 = Decides alone→**qi_06** 2 = Decides jointly or in consultation with someone else →**qi_05A** 3 = Other person decides →**qi_05B**	|____|
qi_05a	**Who do you usually make these decisions with?**	01 = Co-wife	☐

	INTERVIEWER: move to qf_06	02 = Husband	☐
		03 = Mother-in-law	☐
		04 = Father-in-law	☐
		05 = Own mother	☐
		06 = Own father	☐
		07 = Other woman in the household	☐
		08 = Other men in the household	☐
		96 = Other	☐
qi_05b	**Who usually decides for you?** *INTERVIEWER: move to qf_06*	01 = Co-wife	☐
		02 = Husband	☐
		03 = Mother-in-law	☐
		04 = Father-in-law	☐
		05 = Own mother	☐
		06 = Own father	☐
		07 = Other woman in the household	☐
		08 = Other men in the household	☐
		96 = Other	☐
qi_06	**Do you usually have to ask permission from your husband /other influential man in the household to…**		
	A. …go alone to the market?	1 = Yes 2 = No 9 = Never goes to the market	|____|
	B. …go alone to the next village?	1 = Yes 2 = No 9 = Never goes to the next village	|____|
	C. …go alone to the nearest health facility?	1 = Yes 2 = No 9 = Never goes to the health facility	|____|
qi_07	**In your opinion, is a husband/other influential man in the household justified in hitting or beating his wife…**		
	A. …if she goes out without telling him?	1 = Yes 2 = No	|____|
	B. …if she neglects the children?	1 = Yes 2 = No	|____|

	C. …if she argues with him?	1 = Yes 2 = No	|____|
	D. …if she burns the food?	1 = Yes 2 = No	|____|

SECTION J: Infant and Young Child Feeding

Question number	Question and instructions	Answer codes	Answer
Now, we are going to talk about counselling received in the community (A GARI) or in your house. In this counselling at the community women usually are trained or educated on topics related to breastfeeeding, feeding practices and the care of babies and young children			
qj_01	**Have you ever received any specific training on breastfeeding and feeding practices of infants and young children <u>in your community?</u>** **<u>This includes visits at your house, family or community ceremonies, or support groups that meet somewhere in the community?</u>** *<u>INTERVIEWER</u>: Probe: Use* **SHOWCARD # 14** *This question does NOT refer to training on breastfeeding at the health facility, even if the health facility is in the community.*	1 = Yes 2 = No .....→ Error! Reference source not found.	|____|
qj_02	**In which setting(s) have you EVER received this training in your community?** *<u>INTERVIEWER:</u> Do not read the list of answers aloud.* *Please cross the box (☒) for all that apply.* *Keep asking "**anywhere else?**" until the respondent has given all the settings in which she has received training*	01 = At your own house or neighbour's house	☐
		02 = At community volunteer's house	☐
		03 = At family or community ceremony, like naming ceremony or marriage ceremony	☐
		04 = At or nearby the village head/community leader's house	☐
		05 = At a public space in the community, parks, primary schools.	☐
		06 = Other. Specify (_____)	☐
qj_03	**<u>How long ago</u> was the <u>first time</u> you received this type of training in your community or house?** *INTERVIWER: If less than 1 year ago write # of months and '00' in years* *If answer is less than 10 years or less then 10 months, type '0' followed by the number of years/months* *98 years 98 months = Don't know* *The number of months should be lower than 12*		|__|_| years and |__|_| months ago

Question number	Question and instructions	Answer codes	Answer
qj_04	**How long ago was the last time you received this type of training in your community or house?** *INTERVIEWER: if less than 1 year ago write # of months and '00' in years* *If answer is less than 10 years or less than 10 months, type '0' followed by the number of years/months* *The number of months should be lower than 12.* *Number of months and years ago the woman received training for the LAST time should be less or equal to the number of months and years ago the woman received training the FIRST time*	98 years 98 months = Don't know	|__|__| years and |__|__| months ago
qj_05	**How many times have you received this type of training in the last 6 months?**	98 = Don't know	|___|___| times
qj_06	**The last time you received this training in the community or at your house, who gave the training on infant and child feeding?** *INTERVIEWER: Do not read the list of answers aloud.* *Please cross the box (☒) for all that apply.* *Keep asking "**anywhere else?**" until the respondent has given all the people who gave training the last time* *"Don't know" can't be combined with other answers.*	1 = Community volunteer	☐
		2 = medical / health facility staff	☐
		3 = community or religious leader	☐
		4 = another mother	☐
		5 = someone else. Specify (_____)	☐
		98 = Don't know	☐
qj_07	**The last time did you receive the training in a one-to-one counselling or in a group session?** *INTERVIEWER: Select all that apply*	1 = One-to-one counselling → **if only marks this option go to** Error! Reference source not found.	☐
		2 = Group session	☐
qj_08	**The last time, how many people participated in the group training?** *INTERVIEWER: The answer to this question should be greater than 0.*	98 = Do not know	|____|

Question number	Question and instructions	Answer codes	Answer
qj_09	**Have you heard of food demonstration sessions that take place in the community?** **These are trainings where community volunteers showcase how to cook nutritious recipes using local ingredients. They bring prepared or unprepared foods to show to women**	1 = Yes 2 = No ..... → **Next Section** ..... → **Next Section**	|____|
qj_10	**Have you ever attended these food demonstration sessions?**	1 = Yes 2 = No ..... → **Next Section** …….....→ **Next Section**	|____|
qj_11	**How many times have you attended these food demonstration sessions in the LAST 6 MONTHS?**	98 = Don't know	|___| times

SECTION K: Maternal, new-born and child health weeks

Question number	Question and instructions	Answer codes	Answer
Now, I would now like to ask you about <u>EVENTS</u> organised by the State Government programme <u>THAT HAPPEN TWO TIMES A YEAR AT A NEARBY HEALTH FACILITY.</u> **At this health facility, many people come from all over TO RECEIVE IMMUNISATIONS AND MEDICINES that look like this (SHOWCARD # 15).** **The medical staff on this day also give young children VITAMIN A drops (SHOWCARD # 9). We're not talking about polio immunisations where people come house to house. This programme is done at the health facility.**			
qk_01	**Have you <u>HEARD</u> of this maternal and child health week?**	1 = Yes 2 = No ..... → **Section L**	|____|
qk_02	*<u>Built-in skip in CAPI:</u> Do not ask the following question in Jigawa.* **Did you <u>ATTEND</u> a maternal and child health week in the LAST 2 MONTHS?**	1 = Yes ..... → Error! Reference source not found. 2 = No	
qk_03	**The previous maternal and child health week was in [FILL MONTH OF LAST MNCH WEEK].** **Did you <u>ATTEND</u>?** **[To be filled in CAPI: Katsina: December 2015, Zamfara: January 2016, Kebbi: February 2016, Jigawa: May/June 2016]**	1 = Yes ..... → Error! Reference source not found. 2 = No	|____|
qk_04	**Why did you not attend the last maternal and child health week?** *<u>INTERVIEWER</u>: Do not read the list of answers aloud.* *Please cross the box (☒) for all that apply.* *Keep asking "**anything else?**" until the respondent has given you all reasons.*	1 = I did not have time ..... → **Section L** 2 = Too far ..... → **Section L** 3 = Too expensive ..... → **Section L** 4 = Not useful ..... → **Section L** 5 = Did not have permission to go ..... → **Section L** 6 = I did not know about it ..... → **Section L** 7 = Did not have children/ was not pregnant → **Section L** 96 = Other. Specify (_____) → **Section L**	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
MNCH IMPLEMENTATION			

qk_05	**At this maternal and child health week, did any of your children get…?**		
	A. Vitamin A drops? These are different from polio drops that are given house to house *INTERVIEWER: Use* **SHOWCARD #9**	1 = Yes 2 = No 98 = Don't know	⌊_⌋
	B. Deworming pills (Albendazole)? These are pills to get rid of stomach worms and usually come in white colour. *INTERVIEWER: Use* **SHOWCARD #10**	1 = Yes 2 = No 98 = Don't know	⌊_⌋
	C. Mid-upper arm circumference (MUAC) measurement and oedema screening? *INTERVIEWER: Use* **SHOWCARD #12**	1 = Yes 2 = No 98 = Don't know	⌊_⌋
qk_06	**When you attended this maternal and child health week, did any of the children who came with you show signs of diarrhoea?**	1 = Yes 2 = No → Error! Reference source not found. 98 = Don't know → Error! Reference source not found.	⌊_⌋
qk_07	**At this maternal and child health week, did any of your children get…?**		
	D. Oral Rehydration Salts (ORS)? *INTERVIEWER: Use* **SHOWCARD #4**	1 = Yes 2 = No 98 = Don't know	⌊_⌋
	E. Zinc tablets?	1 = Yes 2 = No 98 = Don't know	⌊_⌋
qk_08	**At this maternal and child health week, were you given…?**		
	A. Long Lasting Insecticidal Nets (LLINs)? These are mosquito nets treated with insecticidal to kill mosquitos. *INTERVIEWER: Use* **SHOWCARD #8**	1 = Yes 2 = No 98 = Don't know	⌊_⌋
	B. Tetanus toxoid vaccine? This is an injection given to women of child bearing age or during pregnancy in the upper arm to protect the unborn child from developing tetanus	1 = Yes 2 = No 98 = Don't know	⌊_⌋
qk_09	**When you attended this maternal and child health week, were you pregnant?**	1 = Yes 2 = No → Error! Reference source not found. 98 = Don't know → Error! Reference source not found.	⌊_⌋
qk_10	**At this maternal and child health week, did you get…?**		

	C. Folic acid/ Iron folate (iron with folic acid) supplements? **These are pills given to pregnant women to aid healthy developmental of the baby. It usually comes in yellow colour.** *INTERVIEWER: Use* **SHOWCARD #6**	1 = Yes 2 = No 98 = Don't know	|_|
	D. Iron supplements? These are pills that increase the quantity of blood and give strength. It usually comes in red colour *INTERVIEWER: Use* **SHOWCARD #7**	1 = Yes 2 = No 98 = Don't know	|_|
qk_11	**How much time did you spend to arrive to the facility for this maternal and child health week – on a one way trip?** *INTERVIEWER: This includes all forms of travel (walking, taxi, motorcycle, etc.)*	1 = 0 minutes – less than 30 minutes 2 = 30 minutes - less than 60 minutes 3 = 1 hour - less than 2 hours 4 = 2 hour - less than 5 hours 5 = 5 hours or more 98 = Don't know	|_|
qk_12	**How much money did you spend to arrive to the facility at this maternal and child health week – on a one way trip?** *INTERVIEWER: Please note that the units are in Naira.*	0 = Did not spend anything on transport 98 = Don't know	|_____NGN

SECTION L: Knowledge and attitudes towards infant feeding practices

Question number	Question and instructions	Answer codes	Answer
Now I would like to ask you some questions about nutrition of infants and young children.			
ql_01	**What is the first fluid or solid that should be put in the new-born's mouth?** *INTERVIEWER: Do not read the list of answers aloud.*	1 = Breastmilk 2 = Milk such as tinned, powdered or fresh animal milk (goat, cow) 3 = Infant formula 4 = Plain water 5 = Holy water / Islamic water (zamzam) 6 = Dates or honey 7 = Gripe water 8 = Sugar / glucose water 9 = Juice or juice drinks 10 = Yogurt 11 = Thin porridge 12 = Traditional herbs / tea / infusions 96 = Other (specify) (_____) 98 = I don't know	|__|
ql_02	**After delivery, when do you think is the best time a healthy mother should start breastfeeding her child?** *INTERVIEWER: Read list of answers aloud*	1 = 0 – 1 hour (immediately after birth) 2 = More than 1 hour – 24 hours after birth 3 = More than 1 day after birth 4 = Whenever the baby wants 5 = Whenever the mother is ready 96 = Other (specify) |_____| 98 = Don't know	|__|
ql_03	**The first milk that comes when a mother begins breastfeeding is thick yellowish milk (called colostrum). Do you think this (colostrum) should be given to the baby or should it be discarded?**	1 = Should be given to baby 2 = Should be discarded 98 = Don't know	|__|
ql_04	**Do you think there should be a period of time when a new born should only be fed breastmilk?**	1 = Yes 2 = No → Error! Reference source not found.	

ql_05	**For how long since birth do you think a baby should receive <u>only breast milk and nothing else</u>?** *<u>INTERVIEWER</u>: Write answer in completed months. If less than one month, write 0. This question is about exclusive breastfeeding – no water or other type of milk.*	00 = If less than 1 month 98 = Don't know	|__|__| months
ql_06	**When it is very hot outside, is it ok to give a young baby under 6 months some water to satisfy the baby's thirst?**	1 = Yes 2 = No 98 = Don't know	|__|
ql_07	**Do you think it is ok to give a young baby under 6 months holy water (zamzam)?**	1 = Yes 2 = No 98 = Don't know	|__|
ql_08	**When the baby is young, less than 6 months, is it best to have standard feeding times or is it best to feed the baby whenever he/she wants?**	1 = Best to have standard feeding times 2 = Best to feed the baby whenever he/she wants 98 = Don't know	|__|
ql_09	**At what age of the child do you think a woman should <u>stop breastfeeding</u>?** *<u>INTERVIEWER</u>: <u>W rite answer in completed months. 00 = If less than 1 month 97 = If woman thinks a child should not be breastfed 98 = Don't know</u>* *<u>This is not exclusively breastfeeding. Thus, this asks about time women should continue breastfeeding along with feeding the baby solid/ semi-solid or soft foods.</u>* *<u>The number of months a child should be EXCLUSIVELY breastfed should NOT exceed the number of months a child should be breastfed IN TOTAL</u>*	Write |_|_| months and 00 years if less than 1 year 00 / 00 = If less than 1 month 97 / 97 = If woman thinks a child should not be breastfed 98 / 98 = Do not know	|__|__| years |__|__| months
ql_10	**At what age is the best time to start feeding a baby other foods in addition to breast milk, this includes water, semi-solid, solid and soft foods?** *<u>INTERVIEWER</u>: Write in months if more than 6 months*	1 = Earlier than 6 months 2 = At six months 3 = Later than 6 months. Specify month (_____) 98 = Don't know	|__|
ql_11	**If you needed advice or information relating to child feeding or looking after babies, who would you talk to?** *<u>INTERVIEWER</u>:* *Do not read the list of answers aloud.* *Please cross the box (☒) for all that apply.* *Keep asking "**anyone else?**" until the respondent has given you all people she would talk to.*	01 = Co-wife 02 = Husband 03 = Mother-in-law 04 = Father-in-law 05 = Own mother 06 = Own father 07 = Other woman in the household 08 = Female neighbours or friends	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

		09 = Community Volunteer (CV)	☐
		10 = Medical / health facility staff	☐
		11 = Traditional Birth Attendant (TBA)	☐
		12 = Community or Religious leader	☐
		13 = Nobody	☐
		96 = Other. Specify (_____)	☐
		98 = Don't know	☐

D.3 Referral form

Referral Form (Patient Copy)

Operations Research and Impact Evaluation Project, Endline Survey 2016

Oxford Policy Management

Date: |__|__| / |__|__| / 2016

From: _____

To: _____

Village/LGA: _____

Name of child referred: _____

Age: |__|__| years and |__|__| months Sex: Female ☐ Male ☐

MUAC: |__|__|__| mm Colour of MUAC tape: _____

Weight: |__|__|●|__| kg

Height : |__|__|__|●|__| cm

Referral Centre : _____

Opening days : _____

We kindly thank you of taking the referred child for appropriate follow-up at a health facility.

INTERVIEWER: Look up the closest Referral Centre from the Referral List. Fill in all details and hand to caregiver, head of household or father of the child. Explain her/him that CHILD is showing signs of malnutrition and needs attention by a health professional.

Annex E Key indicators and definitions

The purpose of this annex is to present key definitions of indicators used in this analysis and presented in Volume I and Volume 2.

Table E 1 1. Key indicators list and definitions

Indicator	Numerator	Denominator (Population)	Comment	Source
Dependent variables: IYCF practices indicators				
Proportion of children born ever breastfed	Children aged 0-23 months that were ever breastfed.	All children aged 0-23 months.		WHO (2008, p. 10)
Age appropriate breastfeeding	Infants 0-5 months of age who received only breast milk during the previous day and children 6-23 months of age who received breast milk, as well as solid, semi-solid, or soft foods, during the previous day.	All children aged 0-23 months.		WHO (2008, p.10)
Early initiation of breastfeeding (<1h)	Proportion of children born in the last 24 months who were put to the breast within one hour of birth.	All children aged 0-23 months.		WHO (2008, p. 5)
Early initiation of breastfeeding (<24h)	Proportion of children born in the last 24 months that were put to the breast within 24 hours of birth.	All children aged 0-23 months.		
Exclusive breastfeeding among children < 6 months of age	Infants 0-5 months of age who received only breast milk during the previous day.	All infants aged 0-5 months.	Note that ORS and other medicines are allowed under exclusive breastfeeding. Nothing else is allowed, e.g. no water.	WHO (2008, p. 5)
Continued breastfeeding at 1 year of age (12-15 months of age)	Children 12-15 months of age who received breast milk during the previous day.	All children aged 12-15 months.		WHO (2008, p. 6)
Continued breastfeeding at 2 years of age (20-23 months of age)	Children 20-23 months of age who received breast milk during the previous day.	All children aged 20-23 months.		WHO (2008, p. 10)

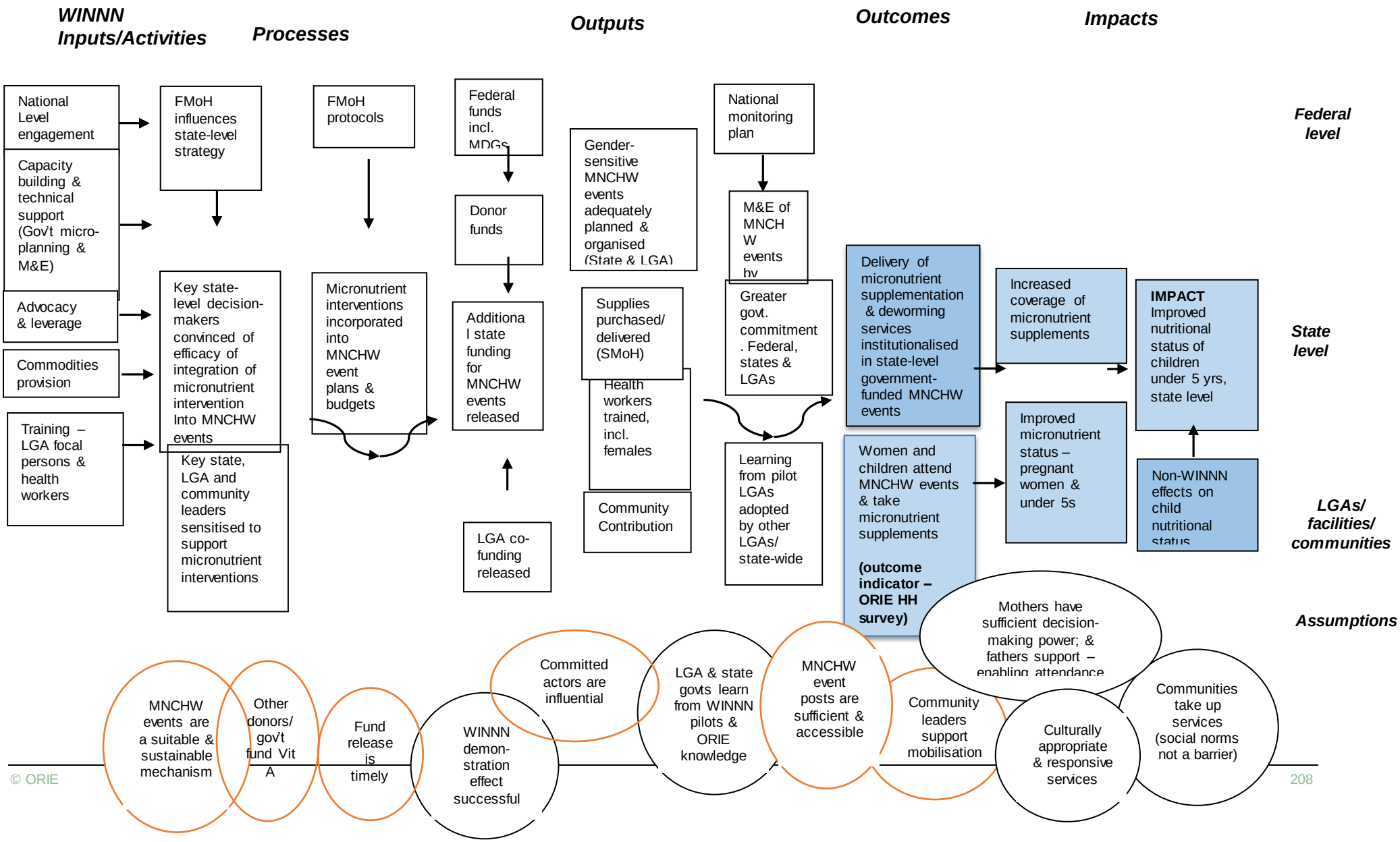
Milk feeding frequency: Proportion of non-breastfed children (6-23 months) who received at least 2 milk feedings during previous day	Currently non-breastfed children 6-23 months of age who received at least 2 milk feedings during the previous day.	All children aged 6-23 months who were currently not breastfed.		WHO (2008, p. 11)
Introduction of solid, semi-solid or soft foods (6-8 months)	Infants 6-8 months of age who received solid, semi-solid, or soft foods during the previous day.	Infants 6-8 months of age.		WHO (2008, p.6)
Consumption of iron-rich or iron-fortified foods (6-23 months of age)	Children 6-23 months of age who received an iron-rich food or a food that was specially designed for infants and young children and was fortified with iron, or a food that was fortified in the home with a product that included iron during the previous day.	All children aged 6-23 months.		WHO (2008, p. 9)
Minimum meal frequency (6-23 months of age)	Breastfed children 6-23 months of age who received solid, semi-solid, or soft foods the minimum number of times or more during the previous day and non-breastfed children aged 6-23 months of age who received solid, semi-solid or soft foods or milk feeds the minimum number of times or more during the previous day.	All children aged 6-23 months.	Minimum is defined as: 2 times for breastfed children aged 6-8 months, 3 times for breastfed children aged 9-23 months, 4 times for non-breastfed children aged 6-23 months.	WHO (2008, p. 8)
Minimum dietary diversity (≥ 4 food groups) (6-23 months of age)	Children 6-23 months of age who received foods from >= 4 food groups during the previous day.	All children aged 6-23 months.		WHO (2008, p. 7)
Minimum acceptable diet (6-23 months of age)	Breastfed children 6-23 months of age who had at least the minimum dietary diversity and the minimum meal frequency during the previous day and non-breastfed children 6-23 months of age who received at least 2 milk feedings and had at least the minimum dietary diversity not including milk feeds and the minimum meal frequency during the previous day.	All children aged 6-23 months.		WHO (2008, p. 8)
Dependent variables: preventive health care				
Percentage of children that are fully vaccinated for age (12-23 months of age)	Children 12-23 months of age who received at least one dose of BCG vaccine, three doses of DPT vaccine, three doses of polio vaccine (excluding polio at birth), and one dose measles vaccine.	All children aged 12-23 months.	Note that in May 2013 pentavalent vaccines were introduced in Nigeria. The present survey accounts for this and counts pentavalent vaccines equivalent to DPT vaccines.	NDHS (2008, p. 145) and NDHS (2013, p. 23)
Dependent variables: child anthropometry				

Percentage of children stunted (0-35 months of age)	Children 0-35 months of age who have a height-for-age z-score below -2 SD of the WHO reference.	All children aged 0-35 months.	The WHO macro available at http://www.who.int/childgrowth/software/en/ was used to calculate this indicator.	WHO (1995, p. 164)
Percentage of children underweight	Children 0-35 months of age who have a weight-for-age z-score below -2 SD of the WHO reference.	All children aged 0-35 months.	The WHO macro available at http://www.who.int/childgrowth/software/en/ was used to calculate this indicator.	WHO (1995, p. 170)
Percentage of children wasted	Children 6-35 months who have a weight-for-height z-score below -2 SD of the WHO reference.	All children aged 6-35 months.	The WHO macro available at http://www.who.int/childgrowth/software/en/ was used to calculate this indicator.	WHO (1995, p. 165)
Percentage of children severely wasted	Children 6-35 months who have a weight-for-height z-score below -3 SD of the WHO reference.	All children aged 6-35 months.	The WHO macro available at http://www.who.int/childgrowth/software/en/ was used to calculate this indicator.	WHO (1995, p. 165)
Percentage of children 6-35 months who are severely acutely malnourished based on MUAC measurement and/or oedema	Children 6-35 months who have a MUAC below 115mm and/or have visible oedemas, that is "swelling caused by the accumulation of fluid in the body tissues" (WHO, 2017).	All children aged 6-35 months.	To asses oedema, the questionnaire specifies the following: "The ANTHRO SPECIALIST should take the following steps: STEP 1: Gently press your thumb into the top of the child's foot for 3 SECONDS STEP 2: Take your finger off the foot and check if there is a deep indentation or 'finger print' If there is a deep indentation or 'finger print' that remains in the top of the child's foot – there is OEDEMA If the top of the foot returns to normal – there is NO OEDEMA".	
Percentage of children 6-35 months who are severely acutely malnourished based on WHZ and/or oedema	Children 6-35 months who have a weight-for-height z-score below -3 SD of the WHO reference and/or have visible oedemas, that is "swelling caused by the accumulation of fluid in the body tissues" (WHO, 2017).	All children aged 6-35 months.	To asses oedema, the questionnaire specifies the following: "The ANTHRO SPECIALIST should take the following steps: STEP 1: Gently press your thumb into the top of the child's foot for 3 SECONDS STEP 2: Take your finger off the foot and check if there is a deep indentation or 'finger print' If there is a deep indentation or 'finger print' that remains in the top of the child's foot – there is OEDEMA If the top of the foot returns to normal – there is NO OEDEMA".	
Dependent variables: WINNN exposure variables				
Proportion of mothers who ever attended IYCF counselling in the community	Mothers 15-49 years who have ever attended IYCF counselling in the community of any type (groups or alone, by community volunteers or through support group)	All mothers 15-49 years	-	
Proportion of mothers who have attended the last MNCH weeks	At baseline: Mothers 15-49 years who have attended the May round (2013) of MNCH weeks, except in Kebbi At endline: Mothers 15-49 years who have attended the May round (2016) of MNCH weeks in Jigawa and either the May (2016) or November(2015) round of MNCH weeks in Katsina, Kebbi and Zamfara.	All mothers 15-49 years	-	

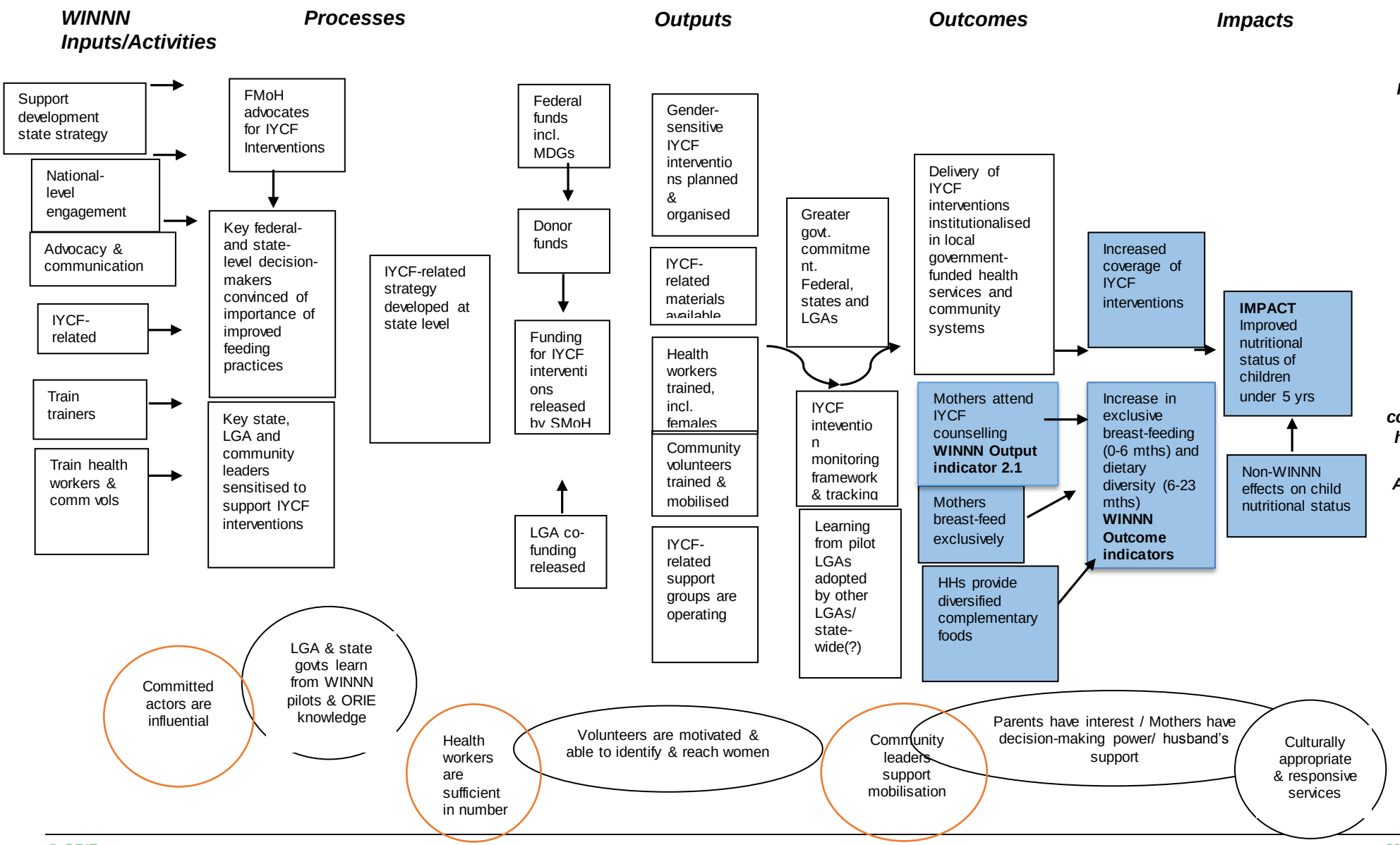
Proportion of children who ever had their MUAC measured	Children 6-35 months who ever had their MUAC measured either at the community, at the health facility or through MNCH weeks.	All children aged 6-35 months.		
Disaggregation variables: additional key indicators				
Household poverty score	Household poverty score developed using information on household size, housing and sanitary situation, asset ownership, and educational attainment of household members.	Indicator calculated for all households.	The household wealth quintiles were derived from this score.	Shiyuan et al. (2008)
Distance to health facility	Distance was defined as geodetic distance, i.e. “the length of the shortest curve between two points along the surface of a mathematical model of the earth”			Vincenty, 1975

Annex F WINNN Theory of change- Graphic Representation

Theory of Change: WINNN Output 1 – Integration of micronutrient interventions into routine primary health services



Theory of Change: WINNN Output 2 – Infant and Young Child Feeding (IYCF) interventions



Theory of Change: WINNN Output 3 – Community Management of Acute Malnutrition (CMAM) programme